

FISHER

A EU 125 ANS CE 17 FÉVRIER

Marc Bourdeau¹

TITLE

[Quasiquicentennial](#) anniversary of R. A. Fisher [1890 – 1962]

RÉSUMÉ

On profite de cet anniversaire, R. A. Fisher est né en effet le 17 février 1890, pour rappeler brièvement le rôle fondateur de ce statisticien-généticien de notre discipline. On placera le lecteur dans le contexte de l'époque de sa jeunesse. Puis on renverra à deux articles de Stigler (2006, 2001) décrivant la transformation de Fisher de biologiste-généticien en statisticien-mathématicien. Nous terminerons par la liste, espérons-la exhaustive, des nombreux concepts (termes) dont Fisher est à l'origine. Nous renvoyons aussi le lecteur intéressé à un [fichier de références bibliographiques](#), sûrement non exhaustif mais suffisant, concernant plus ou moins directement Fisher et ses concepts qui font parfois aujourd'hui controverse. Fisher est encore bien vivant, il sera avec nous longtemps.

Mots-clés : *R. A. Fisher, Karl Pearson, Francis Galton, concepts fishériens.*

ABSTRACT

We take heed from the quasiquicentennial anniversary of R. A. Fisher, born on Feb. 17 1890, to recall the seminal role that this statistician-geneticist has played in the discipline of Statistics. We will give a brief survey of Fisher's early life context; describe, following Stigler (2001, 2006), how he became, a mathematical statistician from a geneticist; and finally will provide the reader with a list, hopefully exhaustive, of terms he has coined. A certain number of references, sufficient but not exhaustive, are also provided through a [hyperlink](#). Fisher will stay with us for a long time.

Keywords : *R. A. Fisher, Karl Pearson, Francis Galton, fisherian concepts.*

1 Introduction – La trame

Un cent-vingt-cinquième anniversaire, ce n'est pas courant qu'on célèbre cela, mais R. A. Fisher [1890 – 1962] en vaut la peine. Il a joué le rôle de père fondateur de la statistique moderne qu'on appelle parfois, à juste titre, la statistique fishérienne.

Nous mettrons en évidence quelques éléments du contexte dans lequel s'est déroulée la vie de Fisher et de ses illustres devanciers que sont Francis Galton et Karl Pearson. Puis nous montrerons à partir des travaux de S.M. Stigler (2006, 2001) comment Fisher, de biologiste devint un statisticien-mathématicien qui a révolutionné la science moderne (Salsburg, 2001).

Nous laisserons donc à d'autres le soin de parler de l'homme et de sa vie & œuvre en général (Droesbeke & Tassi, 1990b ; Salsberg, 2001 ; Senn, 2003, 2012 ; Stigler, 1999). On pourra se reporter à Droesbeke & al. (1990a, 2005) pour avoir une idée du côté caractériel bien avéré du personnage (aussi e.g. Edwards & Bodmer, 2012).² On n'en finirait plus de faire

¹ Professeur associé, École Polytechnique de Montréal, Louis.Marc.Bourdeau@Gmail.com.

² Sans oublier l'autobiographie de G. E. P. Box son gendre, pour certains détails intéressants (Box, 2013), ainsi que la biographie de Jane Box (Box, 1978), la deuxième fille de Fisher.

l'histoire de ses querelles épiques avec des statisticiens de renom dont Karl Pearson, George Udny Yule, Jerzy Neyman.³

Pour se replacer dans le contexte de l'époque des fondements de la statistique moderne, nous devons avant tout mentionner l'implication, qu'on n'évoque pas souvent, des statisticiens anglais de cette époque, tout autant des biologistes et férus de darwinisme, dans la théorie de l'[eugénisme](#)⁴ (du grec *eu*, bien, et *gennân*, engendrer) où les idées de l'amélioration scientifique de la race humaine ont eu une vogue certaine jusqu'au début des années trente du XX^e siècle, et même au-delà. Cette 'science' moderne a vu le jour sous l'impulsion de [Francis Galton](#) [1822 – 1911] (le cousin de [Ch. Darwin](#) [1809 – 1882]).⁵ Elle a impliqué nombre de scientifiques dont Karl Pearson et R. A. Fisher tous [inquiets de la dégénérescence de la race humaine](#). Sous leur impulsion on a vu fleurir des campagnes de stérilisation des criminels, des alcooliques, etc. un peu partout dans les pays développés.

Il faut bien comprendre le lien entre la statistique et l'eugénisme : qui dit eugénisme dit hérédité, et qui dit hérédité dit variation dans le bagage génétique (au moins phénotypique, car les gènes ne furent mis en évidence que bien plus tard, on pourrait parler d'héritage, ou de transmission de l'hérédité) qu'il faut bien savoir mesurer. De là serait née une bonne part de la statistique. Peut-être pas tout à fait là une 'bonne naissance', si l'on ose dire...

On sentira encore plus les préoccupations de ces anglais de l'avenir de la race humaine, quand on se mettra dans l'esprit la fascination de l'Angleterre pour le développement des races [canines](#)⁶ et hippiques, sélections non naturelles s'il en est, qui pourrait servir de métaphore pour les bienfaits attendus de l'eugénisme.

Le XIX^e est le siècle de l'Angleterre. Le soleil ne se couche pas sur l'empire de Victoria [1819 – 1901] qui régna de 1837 à sa mort ; la révolution industrielle (basée sur la science et alliée à la technique) s'est enclenchée en Angleterre à la fin du XVIII^e siècle et il a fallu pas mal de temps avant que les autres pays n'emboîtent le pas. Les grandes universités anglaises, autour de Cambridge et Oxford (ou 'Oxbridge' comme on a dit) ont formé pendant des générations, et ça dure encore, des polymathes, des esprits universels. Galton, Pearson, Fisher, Russell, Whitehead en sont des exemples parmi des centaines. On comprend la fierté (supériorité !) que ressentent les Anglais quand ils considèrent leur culture qui a commencé bien avant le 'grand Will' (William Shakespeare), et leur organisation politique et scientifique, d'une belle stabilité. Voir à cet égard, e.g. Bronowski & Mazlich (1960), *Part III*, qui compare les conditions économico-sociales des quatre révolutions du XVIII^e siècle : anglaise, industrielle, américaine, française.

³ La revue TAS (*The American Statistician*), rapporte plusieurs des éléments de ces 'controverses' scientifiques qui ont émaillé la vie de Fisher. Nous avons recueilli plusieurs références pertinentes à l'œuvre de Fisher et de son héritage scientifique, notamment au sujet des controverses reliées à l'utilisation du paradigme fréquentiel en estimation & test, dans le fichier sous-jacent.

⁴ On peut ajouter que le mouvement a commencé avec la publication de l'œuvre capitale de Darwin, [The Origin of Species by Means of Natural Selection](#) en 1859, suivie en 1871 de [The Descent of Man, and Selection in Relation to Sex](#).

⁵ Galton a souvent été impressionné par les transmissions héréditaires. Il a traité du phénomène du génie et de l'intelligence exceptionnelle qui court dans certaines familles, de la taille des enfants par comparaison avec celle des parents, et inventé le terme de 'régression vers la moyenne' qui est un peu trompeur. C'est ainsi qu'il a été un des inventeurs de la corrélation et donc de la régression linéaire. Le lecteur intéressé par cette histoire pourra cliquer sur le lien suivant : «[Le mètre, les moindres carrés et la régression](#)» (Marc Bourdeau) où on la rapporte.

⁶ Darwin dans *The origin of Species* a parlé de l'origine et de la diversité des chiens. On pourra écouter l'épisode (4min) du balado Nova/Science Now qui en fait état, en [cliquant ici](#), ou [encore ici](#) pour l'écouter directement.

Après le départ à la retraite Karl Pearson (1933), son poste de [Galton Professor](#) à l'UCL ([University College London](#)) fut offert en 1934 à Fisher qui l'accepta (Pearson en était parti...) et devint le directeur du *Galton Laboratory for Eugenics*, avant de devenir *Professor of Genetics* à la *Cambridge University* — notez le changement d'appellation... (cf. Stigler, 1999, ch. 19 pour d'autres détails). On retrouve encore les idées de l'eugénisme dans le terme [darwinisme social](#). Qui a encore court dans bien des milieux prétendument progressistes...⁷

2 De la trame au motif

Nous décrivons ici, à partir de deux articles de Stigler (2005, 2006), la transformation de Fisher vers l'âge de 30 ans, qui était plutôt jusque-là un biologiste généticien, en statisticien mathématicien qui a révolutionné l'histoire de la discipline (Salsburg, 2001).

Fisher avait été un étudiant des plus brillants (il a été boursier du *Gonville and Caius college, Cambridge University*)⁸, c'était un brillant mathématicien mais ses preuves négligentes et trop souvent intuitives⁹ l'ont empêché d'être le premier [Wrangler](#) de sa classe, comme Karl Pearson, son «*arch-enemy*» qui le fut, lui.¹⁰ Il avait été introduit dans les cercles de l'eugénisme et a participé, encore étudiant, comme *steward* au *First International Eugenics Conference* en 1912 à Londres. La statistique et la génétique étaient alors étroitement liés.

Des 95 articles qu'il a publié de 1914 à 1920, 92 l'ont été en eugénisme (la plupart des revues de livres), un autre en statistique sur ses liens avec l'eugénisme, et deux seulement peuvent entrer dans la catégorie de publications sur la statistique.¹¹ De ces derniers, il y a tout de même (Fisher, 1915), la détermination de la loi du coefficient de corrélation... L'autre publication (Fisher, 1920) entamait l'ensemble de ses contributions séminales à la théorie de l'estimation statistique.

Mais dès 1916, il avait essuyé un refus de Karl Pearson pour un article rédigé un peu hâtivement (une seule page), qui lui a demandé des précisions — on dirait aujourd'hui des corrections mineures, mais difficiles... Cet article concernait la méthode d'estimation dite gaussienne, dont il a développé des propriétés d'optimalité quelques années plus tard, grâce à ce refus, et qu'il a appelé (Fisher, 1922) la méthode d'estimation par vraisemblance maximum (EVM). Dans son court papier, il avait apporté des réserves sur des résultats d'un article d'une élève de Pearson sans trop être explicite, qui avait utilisé la méthode de Pearson, la méthode des moments... De ce refus, il avait sans doute pris ombrage, mais il était fort occupé par ailleurs. Comme le note Stigler (2006), il a tout de même gardé la lettre de refus toute sa vie... Ses relations avec Pearson sont toutefois restées assez cordiales ... pendant un certain temps.

⁷ [Herbert Spencer](#) [1820 – 1903], à l'origine du darwinisme social, aurait récrit la devise française : «[Liberty, inequality, survival of the fittest](#)». Compte tenu du retour des idées et politiques de droite un peu partout, aux USA, en France, au Canada, au Québec entre autres, on ne peut que d'être saisi d'un peu d'effroi devant ces perspectives.

⁸ Quiconque veut comprendre le fonctionnement de la [Cambridge University](#), et de ses [colleges](#), dont le [Gonville and Caius College](#) pourra cliquer sur les liens sous-jacents qui précèdent.

⁹ À cet égard, les grands mathématiciens sont souvent des intuitifs, leur preuves sont trop courtes, lacunaires, ils «*volent de cime en cime*» (Nietzsche) et laissent à d'autres le soin de rédiger le détail. C'est le [cas de Fisher](#).

¹⁰ [Third Wrangler](#), mais tout de même...

¹¹ Dans l'ensemble de son œuvre publiée (Fisher R. A., *The collected papers of R. A. Fisher*, 1974), on dénombre une moitié environ en eugénisme/génétique, l'autre moitié en statistique. Beaucoup de contributions majeures de part et d'autre.

Aurait-il oublié cet affront? On peut en douter vu ce qui arriva par la suite.

Par ailleurs, dans un livre publié en 1914, l'astrophysicien Eddington (Eddington, 1914), avait fait quelques affirmations douteuses sur la qualité supérieure de l'estimation d'un écart type par un multiple de la moyenne des valeurs absolues des écarts à la moyenne plutôt que par un multiple de l'écart quadratique moyen à la moyenne, et c'est en 1919 que Fisher, un peu avant de joindre l'équipe du *Rothamsted agricultural research station (RAES)*¹² est tombé, par hasard dit-on, sur ces affirmations. Ses puissantes capacités mathématiques lui ont apporté la preuve que Eddington avait tort (Fisher, 1920; Stigler S. M., 1973), ce qu'il reconnut obligeamment, devant la puissance du calcul. Fisher continua ses recherches et publia, en 1922, un mémoire fort complexe mathématiquement parlant, comportant 60 pages, sur les fondements de la statistique théorique (Fisher, 1922). Cette année là, comme si ce n'était pas assez, il publia, ainsi que le note Stigler (2006), 17 revues de livres sur l'eugénisme... Fisher était devenu un statisticien théorique, mais il n'a jamais cessé d'être un généticien.

La statistique du XX^e siècle s'est construite sur ces fondements. Fisher apportait alors des réponses complètes sur les remarques du refus de Pearson de son article en 1916, qui lui avait demandé des justifications à ce qui n'était alors que des affirmations (intuitions) un peu gratuites. Il n'y a pas fait référence, il est probable que Pearson ne s'en soit pas 'souvenu' non plus.

Fisher a alors comparé la méthode des moments de Pearson et la méthode qu'il a appelée l'estimation par vraisemblance maximum (ci-devant la méthode gaussienne, qui remonte en effet à Gauss lorsqu'ils a déterminé la loi des erreurs, appelée en France la loi de Laplace-Gauss et ailleurs la loi normale, terminologie due à K. Pearson... [Droesbeke & Tassi, 1990b]), et démontré la supériorité de la seconde sur la première, contrairement à ce que Pearson avait affirmé dans ses demandes sur l'article qu'il lui a refusé en 1916. Au passage, il a corrigé une erreur dans les degrés de liberté du test du Khi-carré de Pearson qui s'était trompé.

Tout de même, il fut encore assez imprécis dans certaines de ses affirmations et preuves, ce qui a contribué longtemps, et encore, à créer quelques controverses (Aldrich, 1997; Stigler, 2006).

Douce vengeance peut-être... Quoi qu'il en soit, les relations entre Pearson et lui se sont vite dégradées. Cet article eut une influence majeure sur toute la statistique du XX^e siècle. Neyman & Pearson se sont placés dans le nouveau paradigme pour développer ce qu'on appelle la théorie fréquentielle des tests, mais qui suscite encore aujourd'hui des controverses véhémentes. D'autres voies ont été aussi ouvertes par cet article (Stigler, 2006).

À la station expérimentale de Rothamstead, Fisher allait prendre son envol : il publia, trois ans plus tard, à 35 ans, son œuvre la plus importante, d'une originalité foudroyante, et révolutionner, selon le mot de Salsburg (2001), l'ensemble des sciences: *Statistical methods for research workers* (Fisher, 1925).

Ce livre, publié sans son appareil mathématique, fut rapidement extrêmement bien reçu par l'ensemble de la communauté scientifique, mais ne fit au moment de sa publication l'objet

¹² Pearson avait également refusé un autre article de Fisher... Mais, reconnaissant ses grandes capacités, il lui avait offert en 1919 un poste à son laboratoire de statistique au *UCL (University College London)*, où il régnait en maître. Fisher, qui gagnait péniblement sa vie comme instituteur, mais sentant bien une certaine incompatibilité de caractère entre Pearson et lui, s'est empressé d'accepter... une offre concurrente au *RAES*, qui correspondait plus à ses talents de biologiste et généticien et où il aurait sûrement plus de liberté pour développer ses talents. Comme noté plus haut, il revint à Londres en 1934 au moment de la retraite de Pearson...

d’aucun compte rendu favorable! (Box, 2013, p. 153). Il en est à quelques dizaines d’éditions, il a été traduit dans une dizaines de langues. C’est le plus grand succès de librairie des livres statistiques.

Fisher continua à développer la pratique des plans d’expériences et publia en 1935 *The design of experiments* (Fisher 1935), son *magnum opus*. Tout en continuant à publier des travaux qui ont fait époque en théorie de l’évolution. Ce livre présente (chap. 2) l’exemple célébrissime *The lady tasting tea* (Salsburg, 2001, 2012).

Y eût-il un Prix Nobel scientifique général, Fisher aurait dû en être le premier récipiendaire!

Les EVM, les statistiques exhaustives (*sufficient* en anglais) sont la base de la théorie moderne de l’estimation statistique qui continue à irriguer l’ensemble des sciences. Ce qu’il a mis en place, bien d’autres continuent à en vivre. Fisher est encore bien vivant, il sera avec nous encore longtemps. Parions qu’on célébrera son deux centième avec éclat!

Note – On trouvera en cliquant [ici](#), un certain nombre d’articles, dont ceux cités ici, ainsi que certains autres sur les controverses de la théorie des tests statistiques dont on parle beaucoup actuellement.

3 La statistique et ses mots

Dans cette section, nous donnons dans la langue d’origine, sans plus de façon, les termes statistiques dont Fisher a été l’inventeur (tiré de H. A. David, 1995 & 1998).¹³ On a ici les mots, les choses sont dans les traités de statistique...

TABLEAU 1 – Les termes statistiques introduits par Fisher et leurs dates de première occurrence.

Terme	Première occurrence	Terme	Première occurrence
Variance	1918	Test of significance	1925
Analysis of variance	1918	Level of significance	1925
Likelihood	1921	Sufficient statistic	1925
Maximum likelihood	1922	Randomization	1926
Parameter	1922	Randomized blocks	1926
Statistic	1922	Confounding	1926
Consistency	1922	Interaction	1926
Efficiency	1922	Sampling distribution	1926
Sufficiency	1922	Covariance	1930
Degrees of freedom	1922	Greco-Latin square (with F.Yates)	1934
Student’s t	1924	Null Hypothesis	1935
Z-Distribution	1924	Yates’s correction for continuity	1936
		Normal score (with F. Yates)	1938

¹³ On pourra également se rapporter au site [sous-jacent](#).

Références

- Aldrich, J. (1997). R. A. Fisher and the making of maximum likelihood 1912-1922. *Statistical Science*, 12(3), pp. 162-176.
- Box, G. E. P. (2013). *An accidental statistician*. New York, NY: John Wiley & Sons.
- Box, Jane [née Fisher] (1978). *R. A. Fisher: The life of a scientist*. New York NY: John Wiley & Sons.
- Bronowski, J., & Mazlich, B. (1960). *The western intellectual tradition. From Leonardo to Hegel*. New York, NY: Harper & Row Publishers.
- David, H. A. (1995). First (?) occurrence of common terms in mathematical Statistics. *The American Statistician*, 49(2), pp. 121-133.
- David, H. A. (1998). First (?) occurrence of common terms in mathematical Statistics. A second list with corrections. *The American Statistician*, 52(1), pp. 36-40.
- Droesbeke, J.-J., & Tassi, P. (1990a). George Udny Yule ou comment (ne pas) parler de corrélation. *Statistique et Analyse des données*, 15, pp. 25 - 43.
- Droesbeke, J.-J., & Tassi, P. (1990b). *Histoire de la statistique*. Paris F: Presses Universitaires de France, Coll. Que-sais-je? no. 2527.
- Eddington, A. S. (1914). *Stellar movements and the structure of the universe*. London, UK: Macmillan.
- Fisher, R. A. (1915). Frequency distribution of the value of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10, pp. 507-532.
- Fisher, R. A. (1920). A mathematical examination of the methods of determining the accuracy of an observation by the mean error, and the mean square error. *Monthly Notices of the Royal Astronomical Society*, 80, pp. 758-770.
- Fisher, R. A. (1922). On the mathematical foundation fo theoretical Statistics. *Philosophical Transactions of the Royal Societay of London (A)*, 222, pp. 309-368.
- Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh, UK: Oliver and Boyd.
- Fisher, R. A. (1935). *The design of experiments*. Edinburg, UK: Oliver and Boyd.
- Fisher, R. A. (1974). *The collected papers of R. A. Fisher*. (J. H. Bennet, Éd.) Adelaide, AUS: University of Adelaide Press.
- Heyde, C. C., & Seneta, E. (2001). *Statistician of the centuries*. New York, NY: Springer-Verlag.
- Salsburg, D. (2001). *The lady tasting tea. How Statistics revolutionized science in the twentieth century*. New York, NY: W. H. Freeman & henry Holt and Co.
- Senn, S. (2003). *Dicing with Death. Chance, Risk and Health*. Cambridge UK: Cambridge University Press.
- Senn, S. (2012, December). Tea for three. Of infusions and inferences and milk in first. *Significance*, pp. 30-33.
- Stigler, S. M. (1973). Laplace, Fisher, and the discovery of the concept of sufficiency. *Biometrika*, 60, pp. 439-445.
- Stigler, S. M. (1999). *Statistics on the table. The history of statistical concepts and methods*. Cambridge, MA: Harvard University Press.
- Stigler, S. M. (2005). Fisher in 1921. *Statistical Science*, 20, pp. 32-49.
- Stigler, S. M. (2006). How Ronald Fisher became a mathematical statistician. *Mathématiques et sciences humaines*, 176(4), pp. 23-30.
- Walters, A. E., & Bodmer, W. (2012, December). R. A. Fisher – 50 years on. *Significance*, pp. 27-29.