



**HAL**  
open science

# ”Statistical Methods And Scientific Inference” Sir R. A. Fisher

J. R. Barra

► **To cite this version:**

J. R. Barra. ”Statistical Methods And Scientific Inference” Sir R. A. Fisher. Annales de l’ISUP, 1957, VI (1), pp.3-10. hal-04094495

**HAL Id: hal-04094495**

**<https://hal.science/hal-04094495v1>**

Submitted on 11 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# "STATISTICAL METHODS AND SCIENTIFIC INFERENCE"

SIR R. A. FISHER

Compte-Rendu présenté par J. R. Barra

Examiner les bases logiques de l'analyse statistique et ses modes de raisonnement, tel est le but que se propose Sir R. Fisher dans son dernier livre : "Statistical Methods and Scientific Inference".

Et certes le besoin s'en faisait sentir; si dans l'immense développement des statistiques scientifiques ou appliquées l'auteur note deux époques : d'abord essentiellement avec Pearson et Galton l'étude mathématique des problèmes, puis plus récemment les essais d'expériences planifiées, il n'en affirme pas moins que la construction d'expériences probantes et leur interprétation doivent s'intégrer dans un ensemble logique et cohérent : notre mode d'investigation du monde réel. Il s'agit d'établir les bases même du raisonnement inductif statistique, n'hésitons donc pas à assurer parfaitement nos étapes de raisonnement.

Le plan suivi par Sir R. Fisher est le suivant :

A - LA TENTATIVE DE BAYES des probabilités à priori et les critiques qui lui ont été faites. Pour l'auteur, ces critiques montrent que la même forme de raisonnement ne pourra pas être appliquée dans toutes les situations et il conviendra, suivant les cas, d'utiliser, soit la méthode de Bayes, soit un test de signification, soit ...

B - QUELQUES FORMES GENERALES DE RAISONNEMENT :

Le test d'une hypothèse et la valeur qu'il convient de lui donner, puis la méthode fiduciaire d'estimation.

C - L'étude comparée de la notion de probabilité et du concept de vraisemblance sur des exemples.

D - Problèmes particuliers à l'estimation.

Au cours de son ouvrage, l'auteur fait de nombreuses incidences sur la notion de probabilité, tant pour remarquer la diversité de significations pratiques que chaque esprit lui donne, que pour justifier la méthode des probabilités fiduciaires. Nous analyserons ces remarques dès maintenant, pour apporter quelques éclaircissements à la suite.



## I. - LA NOTION DE PROBABILITÉ

Les points essentiels pour Sir R. Fisher semblent être les suivants :

$\alpha$  - La définition mathématique n'est qu'un ensemble d'axiomes à utiliser quand nécessaire (p. 32) la théorie axiomatique n'a jamais été prise bien au sérieux dans les branches appliquées (p. 107). Le facteur d'incertitude y est inopérant (p. 109).

$\beta$  - Pour appliquer les résultats mathématiques on doit prendre certaines précautions: par exemple, pour qu'un joueur en expectative devant le jet d'un dé, soit obligé d'admettre la probabilité  $1/6$ , il faut et il suffit qu'il n'ait aucune possibilité pour distinguer entre elles les épreuves futures, donc que son ignorance complète soit bien établie (p. 32).

$\gamma$  - En remarquant le rôle de l'ignorance comme donnée, l'auteur est amené à critiquer la conception de Laplace de la probabilité. Renouveau contre lui l'objection de confusion entre possible et probable - et que d'ailleurs tous les auteurs ne trouvent pas aussi importante - Sir R. Fisher estime que la définition Laplacienne de la probabilité est exprimée de telle façon que l'on puisse confondre les deux propositions :

a) J'attriburai la même probabilité à un événement au cours d'épreuves futures, car je ne peux rien retenir pour les distinguer.

b) Je considérerai des hypothèses comme également probables si je ne peux rien retenir pour les distinguer (p. 34).

Quand l'auteur pense à la théorie axiomatique, il semble que ce soit à la théorie des fonctions complètement additives d'ensemble, cependant on ne saurait oublier qu'avant de considérer un événement comme élément d'un ensemble muni d'une loi de probabilité, on a dû, non seulement faire quelques abus de langage du type : confondre probabilité pour que la proposition "à la suite de telle épreuve l'évènement se produira" soit vraie, avec probabilité de l'évènement par rapport à l'épreuve considérée, mais surtout préciser la notion de catégorie d'épreuves, notion essentielle (Fréchet ...) qui doit être mise en évidence dans tous les cas délicats et que l'on peut considérer comme représentative du "facteur d'incertitude".

Quant à la condition ( $\beta$ ) n'est-elle pas aussi nécessaire à la validité de la probabilité mathématique elle-même? N'est-ce pas elle qui justifie l'assertion suivante : Si  $p$  est la probabilité de  $E$  sur la catégorie d'épreuves représentée par  $l$  seul jet,  $p^n$  est la probabilité de  $E$  sur la catégorie d'épreuves représentée par  $n$  jets? Autrement dit l'indépendance des épreuves. Il apparaîtrait donc que les propositions (a) et (b) de ( $\gamma$ ) ne soient pas comparables. (a) justifie l'indépendance des épreuves (b) est relatif à un mode de formation de la loi de probabilité. Et certes en ce qui concerne (b), comme le montre Sir R. Fisher, les critiques de Venn et Boole, doivent servir à éviter un usage abusif de l'ignorance.

Ainsi les points ( $\alpha$ ) ( $\beta$ ) ( $\gamma$ ) prennent-ils toute leur valeur comme précautions à prendre en calcul des probabilités mais ne pourraient infirmer l'image mathématique que l'on peut actuellement construire des problèmes probabilistes! Mais si l'on se préoccupe de rigueur, ne doit-on pas s'astreindre à ne parler de probabilité uniquement pour des éléments aléatoires? Cette idée avait déjà conduit M. Fréchet à proposer probabilités des hypothèses au lieu de probabilités des causes. On pourra constater dans la suite qu'un tel souci évite bien des embarras.



## II. - LE CONCEPT DE VRAISEMBLANCE ET LE PROBLÈME STATISTIQUE

Parallèlement à la notion de probabilité, Sir R. Fisher souligne avec une particulière vigueur l'importance de la notion de vraisemblance et même le rôle complémentaire de celle-ci par rapport à la probabilité dans la représentation mathématique de l'incertitude. Sir R. Fisher avait déjà remarquablement exprimé ceci dans "Inverse probability" dans les termes :

"Connaissant la population nous pouvons exprimer notre connaissance incomplète de l'échantillon à venir en termes de probabilité.

Connaissant l'échantillon nous pouvons exprimer notre connaissance incomplète de la population en termes de vraisemblance".

L'auteur note la différence entre ces deux concepts qui n'obéissent pas aux mêmes lois mais qui pourtant mesurent tous deux une certaine "croyance rationnelle". Déduire de là une schématisation mathématique simple du problème statistique il n'y a qu'un pas, nous le ferons brièvement ne serait-ce que pour préciser et simplifier le langage ; admettons la définition suivante :

Soit un espace fonctionnel  $\mathcal{X}$  de fonctions de probabilité que nous nommerons Hypothèses  $f$  ( $f \in \mathcal{X}$ ). Il s'agit de tirer des conclusions sur  $\mathcal{X}$ , étant donnée une "fonction de vraisemblance sur  $\mathcal{X}$ ", celle-ci définie comme suit

A tout  $f \in \mathcal{X}$  correspond une certaine population, sur toutes ces populations est définie une même catégorie d'épreuves. Supposons que sans que l'on sache sur quelle population on opère on ait appliqué l'épreuve et soit  $\varepsilon$  le résultat. On appellera vraisemblance de  $f$ ,  $L_{\varepsilon}(f)$  : "probabilité pour que l'épreuve étant appliquée à la population correspondant à  $f$  le résultat en soit  $\varepsilon$ ".

On utilise quelquefois des quantités proportionnelles (vraisemblance relative) ou encore le logarithme de  $L$ .

Sir R. Fisher souligne souvent au cours de son ouvrage que la notion de probabilité, ne sera pas toujours susceptible de traduire mathématiquement l'incertitude d'une conclusion. En effet dans la définition précédente, les informations que l'on a sur  $\mathcal{X}$  sont incomplètes et pourtant comment pourrait-on parler de loi de probabilité sur  $\mathcal{X}$  puisque les éléments  $f$  ne sont pas aléatoires? En effet possède-t-on une catégorie d'épreuves sur  $\mathcal{X}$ ?

Notons qu'en pratique :

1° - Les données permettent de réduire  $\mathcal{X}$ , par exemple à un espace fonctionnel isomorphe à  $\mathbb{R}$  (cas paramétrique).

2° - L'épreuve utilisée souvent est un échantillon, mais on utilise ensuite presque toujours une fonction  $T$  (statistique) de cet échantillon et il conviendra, par exemple, de savoir si  $T$  traduit bien les données expérimentales (quantité d'information).

3° - Suivant le but poursuivi on tirera des conclusions orientées dans un certain sens :

- test de signification, c'est-à-dire, choisissant un  $f$  particulier à  $\mathcal{X}$  on étudie la discordance entre l'observation et ce qui aurait été observé si l'épreuve avait été appliquée à la population de  $f$ .

- estimation, choix d'un  $\theta$  optimum, dans quel sens? etc ...



- processus de décision .

Sir R. Fisher examine séparément ces diverses méthodes, sur un grand nombre d'exemples mais dégage particulièrement l'esprit du

### III. - TEST DE SIGNIFICATION

Dans l'analyse que Sir R. Fisher fait de la structure d'un test de signification deux points lui paraissent essentiels.

A - Une vraisemblance très petite jouit de la même valeur psychologique qu'une probabilité très petite.

Supposons que  $L(f_0) - f_0$  hypothèse à rejeter - soit très faible on est placé devant le dilemne suivant (p. 39),

- ou bien un hasard extraordinaire a joué,
- ou bien  $\mathcal{E}$  n'est pas issu de la population correspondant à  $f_0$ .

Sir R. Fisher estime qu'on éprouve la même répugnance à accepter  $f_0$  dans ces conditions qu'à escompter un évènement de probabilité égale à  $L(f_0)$  et déclare (p. 43) :

"Bien qu'identifiable à une condition psychologique de rejet de l'hypothèse, le sentiment introduit par un test de signification a pour base objective un état de probabilité ... et si  $L(f_0)$  représente une mesure des fondements rationnels de refus qu'il engendre il est pourtant évident qu'aucune probabilité n'est établie sur l'hypothèse elle-même".

Ainsi on pourrait étendre aux vraisemblances très voisines de 0, le sens pratique que Borel donnait aux probabilités très voisines de 0.

B - Ne pas confondre, comme le font Neymann et Pearson, test de signification et processus de décision, et ceci pour les raisons suivantes :

1° -  $L(f_0)$  petit entraîne la quasi-certitude de se tromper avec une fréquence bien inférieure à  $L(f_0)$ . En effet si  $\mathcal{E}$  est issu de  $f_0$  on se trompe avec la fréquence  $L(f_0)$  et si  $\mathcal{E}$  n'est pas issu de  $f_0$  on ne se trompe pas. Comme un expérimentateur ne choisit pas  $f_0$  d'une façon quelconque on voit qu'en pratique la fréquence d'évènements est très largement inférieure à  $L(f_0)$ . D'autre part pour un processus de décision, il faudrait connaître les deux erreurs ce qui est inutile dans un test de signification puisque si l'on rejette  $f_0$  on se préoccupe peu de la probabilité de l'accepter à tort. De plus ces erreurs sont subordonnées à l'espace  $\mathcal{X}$  choisi, or celui-ci est le fruit de l'imagination du statisticien et n'est pas comparable à la population bien déterminée et ayant un sens pratique d'un processus classique d'acceptation. Enfin quand le test s'applique à un ensemble d'hypothèses, les calculs précédents ne sont strictement plus possibles. Nous noterons cependant qu'un effort a été fait pour remédier à ceci avec les structures semblables.

2° - C'est enfin méconnaître le caractère d'un test de signification que d'imposer rigidelement une fréquence d'erreur. A la suite d'un test, un expérimentateur ne prend qu'une décision provisoire, jusqu'à plus ample informé, et suivant les cas se montrera plus ou moins strict sur  $L(f_0)$ . Et Sir R. Fisher s'indigne à l'idée qu'un chercheur ne soit qu'un automate sans responsabilité ni initiative au service d'une immense organisation à lois imposées.



Et l'auteur conclut en se défendant d'introduire des "fonctions de prix dans les travaux scientifiques comme on doit le faire dans un processus de décision; car on cherche des méthodes également convaincantes à tous les esprits libres, indépendamment des intentions que l'on pourrait avoir quand à l'utilisation du savoir obtenu".

Le test de signification étant en quelque sorte l'analyse immédiate de la vraisemblance, le deuxième thème de l'ouvrage de Sir R. Fisher est la comparaison entre la méthode fiduciaire et celle des probabilités a priori (Bayes) qui fournissent l'une et l'autre, dans des cas simples, une sorte de fonction de probabilité sur  $\mathcal{X}$ . Est-ce justifié?

#### IV. - MÉTHODE DE BAYES

Sir R. Fisher a déjà exposé cette méthode dans "Inverse probability" :

Supposons que la population dont est tirée l'observation ait été elle-même tirée au hasard d'une super-population de distribution connue a priori d'élément  $\phi(f)$ , la distribution a posteriori est :

$$\frac{\Psi(f) L_{\epsilon}(f)}{\int_{(f)} L_{\epsilon}(f) \phi(f)}$$

De plus en l'absence d'information on prendra  $\phi$  uniforme.

S'attachant plus particulièrement à la validité de cette méthode, Sir R. Fisher à la suite de Venn et Boole, estime que la distribution a priori doit correspondre à une véritable donnée. Exemple (p. 18) :

Soient des souris noires de deux types BB et Bb satisfaisant à

- (I) BB + grise  $\rightarrow$  noire
- (II) Bb + grise  $\rightarrow$  50% de grises et 50 % de noires
- (III) Bb + Bb  $\rightarrow$  25% BB, 50% de Bb et 25% de grises.

Soit une souris noire d'une portée III, avec une grise elle donne 7 noires, est-elle BB ou Bb?

probabilités a priori	BB:1/3	Bb:2/3	
vraisemblance	BB : 1	Bb:1/128	
		$1 \times 1/3$	
probabilités a posteriori	}		$BB = \frac{1 \times 1/3 + 2/3 \times 1/128}{1 \times 1/3 + 2/3 \times 1/128} = \frac{64}{65}$
			$Bb = \frac{2/3 + 1/128}{1 + 1/3 + 2/3 \times 1/128} = \frac{1}{65}$

La donnée : "La souris appartient à une portée III" justifie ici les probabilités a priori de 1/3 et 2/3 mais dire qu'en l'absence d'information elle avait les probabilités a priori de 1/2 et 1/2 reviendrait à traiter le problème comme si elle était issue d'une portée donnant des BB et Bb à 50% et serait donc abusif.

Mais ne faudrait-il pas insister sur le fait que l'esprit accepte une loi de probabilité sur  $\mathcal{X}$  (ici réduit à deux propositions parce que le fait d'être BB est évidemment aléatoire pour une souris noire issue de III, sur l'épreuve facile à concevoir : choix d'une souris noire parmi les noires d'une portée III?



Sir R. Fisher traite des exemples, en particulier prédiction sur de nouvelles épreuves (Règle de Succession) mais souligne qu'on ne doit pas oublier - même si le paramètre a disparu du résultat - que ces déductions supposent la validité des probabilités a priori.

## V. - MÉTHODE FIDUCIAIRE

A la méthode précédente Sir R. Fisher oppose sa méthode fiduciaire (p. 51) qui n'est valable qu'en l'absence complète de distribution a priori mais qui exige :

1° - Des observations "continues", c'est-à-dire pratiquement suffisamment précises pour être considérées comme des valeurs observées d'une variable continue (p. 50).

2° - Il existe une statistique exhaustive, c'est-à-dire :

$$L(f) = L(f)$$

pour  $\varepsilon$  pour T

Supposons encore que T soit estimation de  $\theta$  et que si T est tirée d'une population de paramètre  $\theta$  on ait :

$$P = \text{prob}(T < t) = F(\theta, t)$$

Avec F continue et monotone par rapport à  $\theta$  et t, et exposons avec Sir R. Fisher la méthode sur un exemple.

Soit une source radioactive, émettant des particules à des intervalles de temps x, indépendants, aléatoires, distribuées en  $\theta e^{-\theta x}$  dx,  $\theta$  inconnu. Mesurons n intervalles  $x_i$  et posons :

$$T = \frac{n}{X} \quad X = \sum_{i=1}^n x_i$$

on a

$$L(\theta) = \theta^n \prod_{i=1}^n e^{-\theta x_i} dx_i = \theta^n e^{-\theta X} \prod_{i=1}^n dx_i$$

prenons la vraisemblance relative

$$\ell(\theta) \theta^n e^{-\theta X} \text{ maximum pour } \theta = \frac{X}{n}$$

d'autre part la distribution de X est

$$\theta^n e^{-\theta X} \frac{X^{n-1}}{(n-1)!} dX \quad (1)$$

d'où si l'on pose

$$\chi^2 = 2 \theta X = 2n \frac{\theta}{T}$$

$\chi$  a une distribution d'après (1) qui est simple et même tabulée d'où :

quelque soit P, il existe  $\chi^2(P)$  tel que

$$\text{Prob}[\chi^2 > \chi^2(P)] = P$$

soit

$$\text{Prob}\left[2n \frac{\theta}{T} > \chi^2(P)\right] = P$$

(Remarquons, en aparté, qu'ici  $\chi^2$  est aléatoire sur la catégorie d'épreuves complexe : mesure de n intervalles).



Sir R. Fisher en déduit alors la proposition

$$\text{Prob} \left[ \theta > \frac{T}{2n} \chi^2(P) \right] = P$$

Ce qui représente à son avis une distribution sur  $\theta$ . Comme justification nous nous limiterons aux propres termes de Sir R. Fisher (p. 54).

1° "Dans l'égalité (1) l'ensemble de référence est l'ensemble de toutes les valeurs de T pouvant être obtenues pour une valeur  $\theta$ . Il, (1), a été démontré pour tout  $\theta$  et est ainsi applicable à tous les couples (T,  $\theta$ ) obtenus pour toutes les valeurs de  $\theta$ . Le couple particulier d'un expérimentateur appartenait à cet ensemble élargi, à l'intérieur duquel la proportion des cas satisfaisant l'inégalité

$$\theta > \frac{T}{2n} \chi^2(p) \text{ est égale à } P''.$$

2° - p. 56 "En conséquence il est essentiel d'introduire l'absence d'information a priori comme une donnée distincte pour démontrer complètement l'applicabilité de la méthode fiduciaire de raisonnement aux cas particuliers expérimentaux auxquels elle est appliquée".

Sir R. Fisher note plus loin qu'avec les notations du début  $\frac{\partial F}{\partial T} dT$  est la vraisemblance de  $\theta$  et  $-\frac{\partial F}{\partial A} d\theta$  la "distribution fiduciaire" élémentaire de  $\theta$ ; et enfin précise qu'il ne faut pas confondre les résultats issus de la "distribution fiduciaire" avec les limites de tolérance.

Nous nous limiterons à indiquer les exemples largement traités par Sir R. Fisher en application des deux méthodes précédentes.

- étude d'un "échantillon normal" paramètres inconnus (p. 78) distribution de t et z,
- un cas de régression linéaire (p. 82),
- prédiction fiduciaire et Bayesane (p. 111 à 121) et encore de nombreux autres.

Enfin dans son dernier chapitre Sir R. Fisher traite de la :

## V. - THÉORIE DE L'ESTIMATION

Rappelant ses travaux précédents l'auteur précise quelques notions particulières à cette branche des statistiques : c'est-à-dire comment construire une statistique - définie à partir des observations - qui puisse être considérée comme une estimation satisfaisante. 3 notions essentielles.

### A - CRITERE DE CONSISTANCE

Remarquons d'abord que si T estime  $\theta$ , f(T) doit estimer f( $\theta$ ) d'où rejeter la condition de non-biaisage.

Ensuite si  $T(a_i)$  est une estimation de  $\varphi(A_i)$

$a_i$  observations

$\theta_i$  paramètres

$T(\bar{a}_i) = \varphi(\theta_i)$  définit une estimation consistante.



Sir R. Fisher montre que cette définition est meilleure que la convergence en probabilité de  $T \rightarrow \varphi$  cas si :

$T'_{N_1}$  est une statistique quelconque définie pour  $N_1$  observations,

$T'_N$  est une statistique asymptotiquement consistante, toujours définie.

$$\frac{1}{N} (N_1 T'_{N_1} + (N - N_1) T'_{N - N_1})$$

satisfait à la condition asymptotique mais coïncide avec  $T'_{N_1}$ , pour  $N = N_1$ .

#### B - CRITERE D'EFFICACITE

La première définition donnée fut asymptotique. Soit  $T$  une statistique dont la distribution devient normale si l'effectif  $n$  de l'échantillon tend vers l'infini,  $V(T)$  la variance de  $T$ .

Si  $nV(T)$  a une limite minimum  $T$  est efficace.

Un souci semblable pour des échantillons finis a conduit au critère de suffisance et à utiliser :

#### C - QUANTITE D'INFORMATION

$T$  traduit-il bien toute l'information contenue dans l'échantillon en vue de  $\varphi(\theta_i)$  à estimer ? Cette question est résolue par la quantité d'information de Sir R. Fisher qui ici examine plus particulièrement le problème des statistiques (ou de l'information) subsidiaires et insiste encore sur l'importance de la vraisemblance dans tous ces problèmes.