

Trois lois des erreurs

L'observation des astres est toujours soumise à des erreurs. Comment choisir la valeur « la plus pertinente » parmi plusieurs mesures ? Laplace, mais aussi Legendre et Gauss, échafaudent des théories qui conduisent in fine à la fameuse loi normale.

« Cette hypothèse a été appelée loi des erreurs. Elle ne s'obtient pas par des déductions rigoureuses ; plus d'une démonstration qu'on a voulu en donner est grossière [...]. Tout le monde y croit cependant [...] car les expérimentateurs s'imaginent que c'est un théorème de mathématiques, et les mathématiciens que c'est un fait expérimental. »

Tels sont les mots d'Henri Poincaré dans son cours sur le *Calcul des probabilités* de 1896. Mais qu'est-ce que cette fameuse « loi des erreurs » en laquelle tout le monde croit ? Pour le savoir, interrogeons Pierre-Simon Laplace, qui est revenu sur cette question à plusieurs reprises, notamment en 1774, 1778 et 1810.

Le « milieu de probabilité »

Dans son *Mémoire sur la probabilité des causes par les événements* (1774), Laplace se propose de « déterminer le milieu que l'on doit prendre entre trois observations données d'un même phénomène » (le « milieu » étant ici, pour Laplace, le *milieu de probabilité*, que l'on appelle aujourd'hui la *médiane*). Pour ce faire, il va d'abord déterminer la loi de la courbe ϕ des « probabilités que l'observation s'éloigne de la Vérité ». Or, nous dit-il, « voici les propriétés de



Henri Poincaré (1854—1912).

cette courbe :

1. Elle doit être partagée en deux parties entièrement semblables [...], car il est tout aussi probable que l'observation s'écartera de la vérité à droite comme à gauche [autrement dit, $\varphi(x) = \varphi(-x)$; φ doit être paire] ;
2. Elle doit avoir pour asymptote la ligne [des abscisses], parce que la probabilité que l'observation s'éloigne de la vérité à une distance infinie est évidemment nulle [autrement dit, φ tend vers zéro à l'infini] ;
3. L'aire entière de cette courbe doit être égale à l'unité, puisqu'il est certain que l'observation tombera sur un des points de la droite [des abscisses] [autrement dit, l'intégrale de φ prise sur l'ensemble des nombres réels vaut 1]. »

Laplace poursuit : « Il est nécessaire de connaître $\varphi(x)$: mais, dans le nombre infini de fonctions possibles, laquelle choisirons-nous de préférence ? Les considérations suivantes peuvent nous déterminer dans ce choix. »

Il commence par rejeter une solution constante, car « si l'on supposait exister un très grand nombre d'observations du phénomène, il est à présumer qu'elles deviendraient d'autant plus rares qu'elles s'éloigneraient de la vérité ».

Ainsi, $\varphi(x)$ doit décroître pour $x > 0$, ainsi que sa dérivée (ce que Laplace nomme « les différences »). Mais « comme nous n'avons aucune raison de supposer une autre loi aux ordonnées qu'à leurs différences », il pose que φ doit être proportionnelle à sa dérivée :

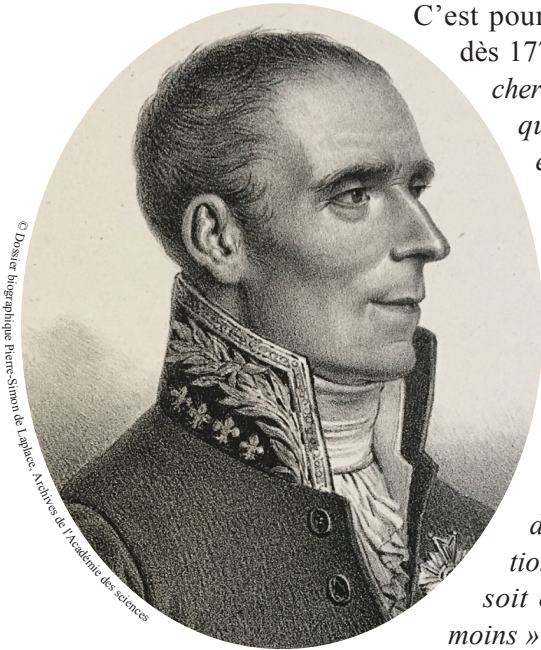
$$\frac{d\varphi(x)}{dx} = -m\varphi(x) \quad \text{pour } x > 0,$$

où la constante $-m$ est négative, d'où (avec les conditions de parité et de normalisation) la première loi des erreurs de Laplace :

$$\varphi(x) = \frac{m}{2} \exp(-m|x|).$$

Aujourd'hui appelée *laplacienne*, cette loi exponentielle est très utile pour modéliser certaines statistiques en traitement du signal.

Laplace n'est, semble-t-il, nullement gêné par la singularité en 0 de sa loi (la dérivée y est discontinue). Mais il n'est pas satisfait pour une autre raison : « On peut objecter contre cette loi qu'en supposant x extrêmement grand $\varphi(x)$ ne serait pas nul, ce qui répugne. »



© Dossier biographique Pierre-Simon de Laplace. Archives de l'Académie des sciences

C'est pourquoi il revient, dès 1777, dans ses *Recherches sur le milieu qu'il faut choisir entre les résultats de plusieurs observations*, sur une nouvelle loi des erreurs qu'il contraint à s'annuler pour $|x| > a$ où a est « la plus grande erreur dont son observation est susceptible, soit en plus, soit en moins ».

Au prix de calculs élaborés sur la courbe moyenne vérifiant les conditions vues précédemment, il trouve une loi logarithmique :

$$\varphi(x) = \frac{1}{2a} \log \frac{a}{|x|},$$

avec une singularité infinie en 0 : c'est la *deuxième loi des erreurs* de Laplace. Celle-ci n'a pas eu le succès de la première, et encore moins de la troisième loi, celle dont parle Poincaré.

Une question de paternité

En 1805, Adrien-Marie Legendre, en appendice de ses *Nouvelles Méthodes pour la détermination des orbites des comètes*, expose la fameuse *méthode des moindres carrés*. Il n'y a aucune mention d'une quelconque « loi des erreurs » chez Legendre ; il ne s'agit que d'une technique simple de calcul, où l'on minimise une somme des carrés des écarts à plusieurs observations pour trouver les valeurs « les plus proches » de la vérité (et donc « les plus vraisemblables »).

Dans le cas d'une seule variable dont on effectue directement plusieurs mesures,

la méthode conduit à faire leur moyenne empirique, que Laplace nomme « *milieu arithmétique* » (et non le « *milieu de probabilité* » qu'il préconisait jusque-là). Une polémique s'installe concernant cette technique, car Gauss s'en attribue la paternité, affirmant l'avoir utilisée dès l'année 1795 (à l'âge de 18 ans). Laplace commentera : « *Ce savant géomètre [M. Legendre] est le premier qui ait publié cette méthode ; mais on doit à M. Gauss la justice d'observer qu'il avait eu, plusieurs années avant cette publication, la même idée dont il faisait un usage habituel, et qu'il avait communiquée à plusieurs astronomes.* »

En 1809, dans sa *Théorie du mouvement des corps célestes*, Gauss justifie la méthode par une loi des erreurs du type $\varphi(x) = \frac{h}{\sqrt{\pi}} \exp(-h^2 x^2)$, la seule

qui conduit aux moindres carrés pour des mesures indépendantes, comme il le démontre par la résolution d'une équation différentielle maximisant la vraisemblance. Cette loi est en conséquence appelée *loi gaussienne* aujourd'hui. Mais est-ce vraiment pertinent de la nommer ainsi plutôt que *troisième loi des erreurs* de Laplace ?

En effet, c'est bien Laplace qui, comme l'écrit d'ailleurs Gauss lui-même, avait déterminé dès 1774, dans son *Mémoire sur la probabilité des causes par les événements*, la valeur de l'intégrale dite de Gauss, que l'on devrait plutôt appeler *intégrale de Laplace* :

$$\int_0^{+\infty} \exp(-t^2) dt = \frac{1}{2} \sqrt{\pi},$$

qui permet de justifier que la « gaussienne » φ vérifie bien la normalisation $\int_{-\infty}^{+\infty} \varphi(x) dx = 1$.

Il aurait été difficile à Gauss, aussi précoce soit-il, de s'attribuer la paternité

de cette intégrale alors qu'il n'était pas encore né !

En outre, l'argument de Gauss est circulaire : il justifie la méthode des moindres carrés par la loi des erreurs « gaussienne », qu'il justifie à son tour par la méthode des moindres carrés ! Rien n'autorise à en déduire quoi que ce soit sur la validité de l'ensemble. C'est ce que Laplace exprime dans sa *Théorie analytique des probabilités* en 1812 : « Comme rien ne prouve que la première de ces règles donne le résultat le plus avantageux, la même incertitude existe par rapport à la seconde. »

C'est finalement Laplace qui, dans son appendice *Mémoire sur les approximations des formules qui sont fonctions de très grands nombres et sur leur application aux probabilités*, trouve dès 1810 la meilleure justification de la loi « gaussienne » : pour un « grand » nombre d'observations, la somme cumulée de petites erreurs indépendantes finit par obéir à cette loi, indépendamment de la loi initiale des erreurs individuelles. Cette justification n'est autre que le *théorème de Laplace*, qui généralise le théorème de Moivre et qu'on appelle aujourd'hui *théorème central limite* (*Zentralen Grenzwertsatz* en allemand, expression introduite par George Pólya en 1920 : c'est bien le théorème qui est central, pas la limite !).

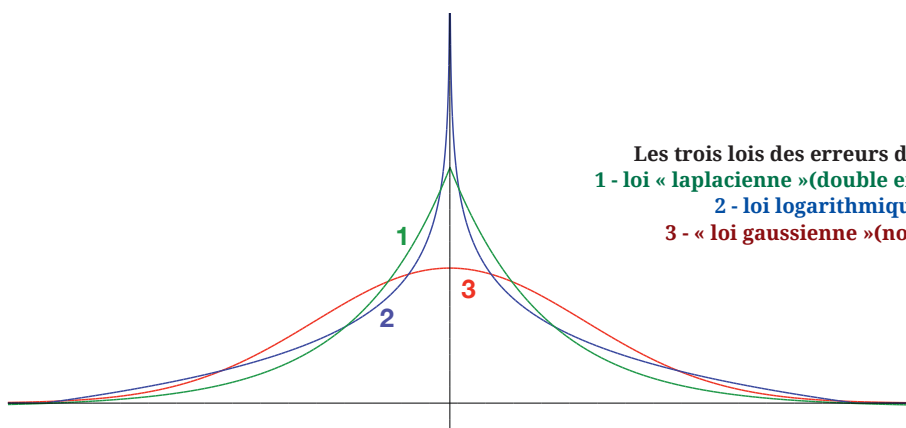
On a eu ensuite du mal à se mettre d'accord sur le nom de cette loi des erreurs : les auteurs français disaient la *loi de Laplace* alors que les auteurs allemands et britanniques parlaient de la *loi de Gauss*. Vers 1950, la Commission de terminologie statistique de l'Association française de normalisation (Afnor), emmenée par le mathématicien Maurice Fréchet, décida de normaliser la locution « loi de Laplace », « *considérant que Laplace a découvert la loi qui devrait porter son nom [...] à une époque où Gauss était encore un jeune enfant ; [la Commission] a rétabli la vérité en rendant l'hommage qui lui était dû* ». Mais malgré ces efforts, le nom de Gauss est resté plus que les autres !

C'est précisément pour ne pas fixer de paternité que de nombreux auteurs ont finalement opté pour le terme *normal*, et en premier Poincaré dans son *Calcul des probabilités* au sujet de l'intégrale

$$\int_{x_0}^{x_1} \frac{h}{\sqrt{\pi}} \exp(-h^2 x^2) dx :$$

« Je dirai, pour abrégé, que la loi de probabilité est normale lorsque la valeur de la probabilité est représentée par cette intégrale. » C'est en effet cette loi qui apparaît le plus souvent dans la Nature et qui se trouve être, généralement, la plus adaptée aux observations...

O. R.



Les trois lois des erreurs de Laplace :
1 - loi « laplacienne » (double exponentielle) ;
2 - loi logarithmique ;
3 - « loi gaussienne » (normale).