LA MÉTHODE DES MOINDRES CARRÉS

François MILLET
Lycée La Prat's, Cluny

Résumé

Les scientifiques cherchent le plus souvent à adapter un modèle mathématique aux phénomènes qu'ils étudient. Malheureusement les données expérimentales ne concordent pas toujours avec le modèle à cause de la présence de différentes erreurs lors des mesures. Après avoir catalogué ces erreurs, il s'est agi de trouver une méthode permettant d'en minimaliser les conséquences. C'est ainsi qu'est née la méthode des moindres carrés par le travail de deux grands savants, Legendre et Gauss, qui en ont chacun revendiqué la paternité. Un exemple de son utilisation par Legendre vient compléter l'exposé.

INTRODUCTION

À partir de la fin du XVIII^{e} siècle, ceux qui se sont intéressés aux statistiques, que ce soit Carl Friedrich Gauss[1], Adrien-Marie Legendre[2], Pierre-Simon (de) Laplace[3], Adolphe Quételet[4], ont été confrontés au problème des erreurs. Ils ont commencé par les classer en fonction de leurs causes et sont arrivés à les ranger dans deux grandes catégories : les erreurs régulières et les erreurs irrégulières. Autant les erreurs régulières étaient assez faciles à gérer, comme nous l'exposerons ensuite, autant les irrégulières ont conduit à des recherches approfondies car il était important d'en minimiser les conséquences.

Carl Friedrich GAUSS (1777-1855)

Ces recherches ont débouché sur la méthode dite « des moindres carrés » qui, après une polémique sur la paternité de celle-ci entre Gauss et Legendre, a été adoptée par la majorité des statisticiens de l'époque et aussi par leurs successeurs jusqu'à nos jours, bien qu'aujourd'hui on utilise plutôt la méthode dite « du maximum de vraisemblance » qui a un champ d'application plus vaste.

---

1 GAUSS, Carl Friedrich, 1777 - 1855. Mathématicien, astronome et physicien allemand. On ne compte plus les domaines auxquels il a apporté sa contribution : nombres premiers, congruences, probabilités, mouvement des corps célestes, électromagnétisme,…

2 LEGENDRE, Adrien-Marie, 1752 - 1833. Mathématicien français, il travailla dans divers domaines tels que la théorie des nombres, les algèbres abstraites, les fonctions elliptiques, les statistiques, la mécanique. C'est dans son ouvrage de mécanique céleste qu'il publie, en 1805, la méthode des moindres carrés.

3 LAPLACE, Pierre-Simon, 1749 - 1827. Mathématicien, astronome et physicien français. Il s'intéressa, entre autres, à la théorie des probabilités. Il démontra, pour la première fois, en 1809 le théorème central limite.

4 QUÉTELET, Adolphe,1796 - 1874. Mathématicien, astronome, naturaliste belge. Il est connu pour ses travaux en statistiques : il est un précurseur de l'étude démographique et le concepteur de « l'homme moyen ».

DES ERREURS

Dans *Théorie de la combinaison des observations qui expose aux moindres erreurs*, présenté à la Société Royale à Göttingen le 15 février 1821, Gauss rend compte de l'omniprésence des erreurs et de la nécessité de les classifier[5] :

> *Quelque soin que l'on apporte aux observations qui concernent la mesure des grandeurs physiques, elles sont forcément soumises à des erreurs plus ou moins considérables. Ces erreurs, dans le plus grand nombre des cas, ne sont pas simples, mais découlent à la fois de plusieurs sources distinctes qu'il est bon de distinguer en deux classes.*

Il précise alors cette classification[6] :

> *Certaines causes d'erreurs dépendent, pour chaque observation, de circonstances variables et indépendantes du résultat que l'on obtient : les erreurs qui en proviennent sont nommées irrégulières ou fortuites, et de même que les circonstances qui les produisent, leur valeur n'est pas susceptible d'être soumise au calcul. Telles sont les erreurs qui naissent de l'imperfection de nos organes et toutes celles qui sont dues à des causes irrégulières extérieures, comme, par exemple, les trépidations de l'air qui rendent la vision moins nette ; quelques-unes de ces erreurs dues à l'imperfection inévitable de nos meilleurs instruments appartiennent à la même catégorie. Nous citerons, par exemple, la rugosité de la partie intérieure du niveau, le défaut de rigidité absolue, etc.*
>
> *Il existe, au contraire, d'autres causes qui, dans toutes les observations de même nature, produisent une erreur identique, ou dépendent de circonstances essentiellement liées au résultat de l'observation. Nous appellerons les erreurs de cette catégorie, des erreurs constantes ou régulières.*

Il admet alors que cette classification n'est pas aussi simple qu'elle y paraît et il poursuit[7] :

> *Il est du reste évident que cette distinction est jusqu'à un certain point relative et dépend du sens plus ou moins large que l'on veut attacher à l'idée d'observations de même nature. Par exemple, si l'on répète indéfiniment la mesure d'un même angle, les erreurs provenant d'une division imparfaite du limbe appartiendront à la classe des erreurs constantes. Si, au contraire, on mesure successivement plusieurs angles différents, les erreurs dues à l'imperfection de la division seront regardées comme fortuites tant que l'on n'aura pas formé la table des erreurs relatives à chaque division.*

Il écarte rapidement les erreurs régulières pour se concentrer sur les irrégulières et axer ses recherches sur les moyens d'en réduire au maximum leur influence[8] :

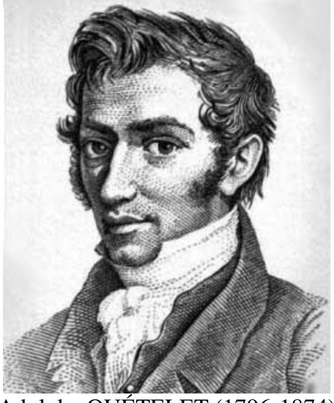
---

[5] GAUSS, Carl Friedrich, *Méthode des moindres carrés*, traduit par J. Bertrand, 1855, page 1.
[6] GAUSS, Carl Friedrich, *Ibid.,* pages 1-2.
[7] GAUSS, Carl Friedrich, *Ibid.*, page 2.
[8] GAUSS, Carl Friedrich, *Ibid.*, page 2.

Nous excluons de nos recherches la considération des erreurs régulières. C'est à l'observateur qu'il appartient de rechercher avec soin les causes qui peuvent produire une erreur constante, pour les écarter si possible, ou tout au moins apprécier leur effet, afin de le corriger sur chaque observation, qui donnera alors le même résultat que si la cause constante n'avait pas existé. Il en est tout autrement des erreurs irrégulières : celles-là, par leur nature, se refusent à tout calcul, et il faut bien les tolérer dans les observations. On peut cependant, par une combinaison habile des résultats, réduire autant que possible leur influence. C'est à cette question importante que sont consacrées les recherches suivantes.

On retrouve une démarche similaire chez Quételet. Dans *Recherches statistiques*[9], présenté le 10 mars 1844 à la Commission centrale de statistique de Belgique dont il est le Président, il fait un exposé sur l'appréciation des documents statistiques. Dans la première partie, intitulée *Appréciation générale des causes et de leurs tendances*, il commence par une critique des méthodes parfois employées par certains[10] :

La statistique compte, chaque jour, plus de prosélytes, parce que chaque jour on en sent mieux le besoin. Toutefois l'espèce d'engouement dont bien des personnes se sont prises pour elle, a plutôt retardé qu'accéléré sa marche. On en a fait d'étranges abus, soit par ignorance, et en violant les premiers principes des sciences d'observation, soit encore par des idées préconçues, et en forçant les chiffres ou en les choisissant de manière à les rendre favorables à des opinions qu'on cherchait à faire prévaloir.


Adolphe QUÉTELET (1796-1874)

Il pense qu'il faut considérer les statistiques comme une science. De plus il affirme que l'utilisation de la valeur moyenne peut engendrer des erreurs importantes et qu'il faut donc se pencher de manière approfondie sur les erreurs et les causes de celles-ci[11].

En se bornant à regarder la statistique comme un art, il suffirait de réunir et de grouper les faits dans un ordre plus ou moins rationnel, plus ou moins facile à saisir ; mais une saine critique doit présider au choix des matériaux, et les règles de la plus sévère logique doivent en diriger l'emploi : c'est là que commence la science.

[…]

La considération des moyennes sert donc en quelque sorte de base à toute la statistique ; mais elle peut donner lieu à de graves erreurs, si l'on n'en fait usage avec circonspection.

[…]

---

9 QUÉTELET, Adolphe, *Recherches statistiques*, 1844.
10 QUÉTELET, Adolphe, *Ibid.*, page 1.
11 QUÉTELET, Adolphe, *Ibid.*, pages 2-3.

Sans recourir à la science, l'habitude nous donne une appréciation vague de la moyenne et des limites des variations qui appartiennent à chaque élément statistique ; c'est d'après cette appréciation que nous sommes guidés dans nos raisonnements. Mais il convient au progrès des lumières de chercher à substituer des idées précises à des notions vagues.

Je ferai remarquer d'abord que les nombres peuvent varier par une infinité de causes, que je partage en trois classes principales :

I. Causes constantes ;

II. Causes variables ;

III. Causes accidentelles.

Les causes constantes sont celles qui agissent, d'une manière continue, avec la même intensité et dans le même sens.

Les causes variables agissent, d'une manière continue, avec des énergies et des tendances qui changent, soit d'après des lois déterminées, soit sans aucune loi apparente. Parmi les causes variables, il importe surtout de remarquer celles qui ont un caractère de périodicité, comme les saisons.

Les causes accidentelles ne se manifestent que fortuitement, et agissent indifféremment dans l'un ou l'autre sens.

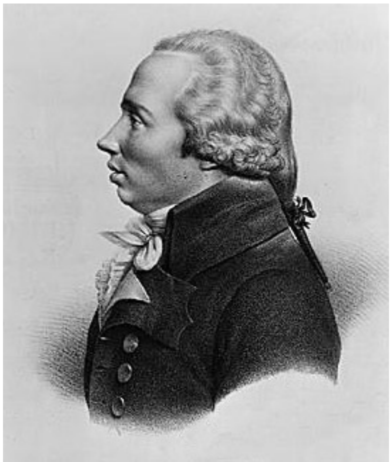
L'emploi des moyennes a surtout pour objet d'éliminer, dans une série de phénomènes observés, les effets des causes accidentelles, et d'arriver à la connaissance des causes constantes et des causes variables.

Or, pour que les statistiques remplissent leur but, elles doivent offrir les moyens de constater ces dernières causes et d'en mesurer le degré d'énergie.

Il s'agit donc des erreurs irrégulières, ou fortuites, qui vont faire l'objet de la méthode des moindres carrés, dans le but d'en minimiser les conséquences.

DEUX APPROCHES DIFFÉRENTES

Gauss et Legendre ont eu, pour établir « leur » méthode des moindres carrés, deux approches différentes. Legendre présente une méthode algébrique alors que Gauss a une approche probabiliste du problème.



Adrien-Marie LEGENDRE (1752-1833)

MÉTHODE DE LEGENDRE

Legendre utilise des équations d'erreurs de la forme $E = a + bx + cy + fz + etc$ dans lesquelles a, b, c, f…sont des coefficients connus et x, y, z sont des variables qu'il faut déterminer de manière à ce que E se réduise pour chaque équation à une quantité nulle ou très petite. Il précise alors[12] :

[12] LEGENDRE, Adrien-Marie, *Nouvelles méthodes pour la détermination des orbites* …, 1805, page 72.

Si l'on a autant d'équations que d'inconnues x, y, z, etc, il n'y a aucune difficulté pour la détermination de ces inconnues, et on peut rendre les erreurs absolument nulles. Mais le plus souvent, le nombre des équations est supérieur à celui des inconnues, et il est impossible d'anéantir toutes les erreurs.

Dans cette circonstance, qui est celle de la plupart des problèmes physiques et astronomiques, où l'on cherche à déterminer quelques éléments importants, il entre nécessairement de l'arbitraire dans la distribution des erreurs, et on ne doit pas s'attendre que toutes les hypothèses conduiront exactement aux mêmes résultats ; mais il faut surtout faire en sorte que les erreurs extrêmes, sans avoir égard à leurs signes, soient renfermées dans les limites les plus étroites qu'il soit possible.

De tous les principes qu'on peut proposer pour cet objet, je pense qu'il n'en est pas de plus général, de plus exact, ni d'une application plus facile que celui dont nous avons fait usage dans les recherches précédentes, et qui consiste à rendre minimum la somme des carrés des erreurs. Par ce moyen, il s'établit entre les erreurs une sorte d'équilibre qui empêchant les extrêmes de prévaloir, est très-propre à faire connaître l'état du système le plus proche de la vérité.

Il considère plusieurs équations $E,\ E',\ E'',\ \ldots$ telles que :

$$E = a + b\,x + c\,y + f\,z + \ldots$$
$$E' = a' + b'\,x + c'\,y + f'\,z + \ldots$$
$$E'' = a'' + b''\,x + c''\,y + f''\,z + \ldots$$

La somme des carrés des erreurs est $E^2 + E'^2 + E''^2 + \ldots$ soit :

$$\begin{aligned} &\left(a + b\,x + c\,y + f\,z + \ldots\right)^2 \\ &+\left(a' + b'\,x + c'\,y + f'\,z + \ldots\right)^2 \\ &+\left(a'' + b''\,x + c''\,y + f''\,z + \ldots\right)^2 \\ &+\ldots \end{aligned}$$

Il faut minimiser cette quantité pour chacune des variables. Cela revient donc à annuler les dérivées partielles. En considérant *x* constante, Legendre obtient ce qu'il écrit sous la forme suivante : $0 = \int a\,b + x \int b^2 + y \int bc + z \int b\,f + \ldots$ et qui correspond à l'égalité $0 = (a\,b + a'\,b' + a''\,b'' + \ldots) + x\,(b^2 + b'^2 + b''^2 + \ldots) + y\,(b\,c + b'\,c' + b''\,c'' + \ldots) + \ldots$ Il détermine de même le minimum par rapport à *y*, ce qui donne $0 = \int a\,c + x \int b\,c + y \int c^2 + z \int f\,c + \ldots$, puis par rapport à *z* soit $0 = \int a\,f + x \int b\,f + y \int c\,f + z \int f^2 + \ldots$ Il fait alors remarquer que l'on retrouve les mêmes coefficients, ce qui facilite les calculs. Il donne donc une méthode[13] :

En général, pour former l'équation du minimum *par rapport à l'une des inconnues, il faut multiplier tous les termes de chaque équation proposée par le coefficient de l'inconnue dans cette équation, pris avec son signe, et faire une somme de tous ces produits.*

---

[13] LEGENDRE, Adrien-Marie, *Ibid.*, page 73.

On obtiendra de cette manière autant d'équations du minimum, qu'il y a d'inconnues, et il faudra résoudre ces équations par les méthodes ordinaires.

Legendre fait alors quelques remarques sur les résultats obtenus, notamment une mettant en avant la simplicité de sa méthode en cas de nouveaux calculs[14].

Si après avoir déterminé toutes les inconnues x, y, z,…, on substitue leurs valeurs dans les équations proposées, on connaîtra les diverses erreurs $E, E', E'', \ldots$ *auxquelles ce système donne lieu, et qui ne peuvent être réduites sans augmenter la somme de leurs carrés. Si parmi ces erreurs il s'en trouve que l'on juge trop grandes pour être admises, alors on rejettera les équations qui ont produit ces erreurs, comme venant d'expériences trop défectueuses, et on déterminera les inconnues par le moyen des équations restantes, qui alors donneront des erreurs beaucoup moindres. Et il est à observer qu'on ne sera pas obligé alors de recommencer tous les calculs, car comme les équations du* minimum *se forment par l'addition des produits faits dans chacune des équations proposées, il suffira d'écarter de l'addition les produits donnés par les équations qui auront conduit à des erreurs trop considérables.*

Il explique ensuite que la méthode de la moyenne utilisée habituellement n'est qu'une conséquence de la méthode des moindres carrés[15].

La règle par laquelle on prend le milieu entre les résultats de différentes observations n'est qu'une conséquence très simple de notre méthode générale, que nous appellerons Méthode des moindres quarrés.

En effet, si l'expérience a donné diverses valeurs a', a'', a''' …pour une certaine quantité x, la somme des carrés des erreurs sera $(a'-x)^2+(a''-x)^2+(a'''-x)^2+\ldots$*, et en égalant cette somme à un minimum, on a* $0=(a'-x)+(a''-x)+(a'''-x)+\ldots$ *d'où résulte* $x=\dfrac{a'+a''+a'''+\ldots}{n}$*, n étant le nombre d'observations.*

Il termine son exposé en revenant sur la simplicité et la portée de cette méthode, propriétés qui seront illustrées par une application géodésique[16].

L'application que nous allons faire de cette méthode à la mesure de la méridienne, achèvera de mettre dans tout son jour sa simplicité et sa fécondité.

MÉTHODE DE GAUSS

Gauss part de *m* fonctions notées V, V', V'', etc., de *n* inconnues notées p, q, r, s, etc., et il suppose que les observations ont donné les valeurs M, M', M'', etc., telles que $V = M$, $V' = M'$, $V'' = M''$, etc...

[14] LEGENDRE, Adrien-Marie, *Ibid.*, page 74.
[15] LEGENDRE, Adrien-Marie, *Ibid.*, pages 74-75.
[16] LEGENDRE, Adrien-Marie, *Ibid.*, page 75.

On obtient donc ainsi, en général, un système de m équations à n inconnues. Gauss ne considère que le cas $m > n$, système plus que déterminé[17].

> *En général, le calcul de ces inconnues constituera un problème indéterminé, déterminé ou plus que déterminé, suivant qu'on aura* $m < n$, $m = n$ ou $m > n$ *.Nous ne nous occuperons ici que du dernier cas, dans lequel évidemment il ne serait possible d'obtenir une représentation exacte de toutes les observations, que si ces observations n'étaient affectées d'aucune erreur. Mais comme cela n'a jamais lieu dans la nature, on devra regarder comme possible tout système de valeurs des inconnues* p, q, r, s, *etc., desquelles résultent, pour les fonctions* $V-M, V'-M', V''-M''$, *des valeurs qui ne surpassent pas les limites des erreurs que l'on peut commettre dans les observations, mais on ne doit pas regarder tous ces systèmes possibles comme jouissant du même degré de probabilité.*

Il va alors s'intéresser à la fonction, notée φ, qui correspond à la probabilité qu'une erreur, notée Δ soit commise. Une partie des ses travaux consiste à déterminer cette fonction ou tout du moins à déterminer une fonction qui l'approche le plus exactement possible et avec laquelle on puisse faire des calculs. Gauss est en effet conscient que la fonction φ ne peut être connue de manière précise[18].

> *La probabilité qu'une erreur Δ soit commise dans l'une des observations sera une fonction de Δ, que nous nommerons φ(Δ). Quoique cette fonction ne puisse être assignée d'une manière précise, on peut du moins affirmer qu'elle doit devenir maximum pour Δ = 0, avoir dans la plupart des cas la même valeur pour des valeurs de Δ égales et de signes contraires, et, enfin, s'évanouir quand on donne à Δ une valeur égale ou supérieure à l'erreur maximum ; φ(Δ) doit donc, à proprement parler, être rapportée à la classe des fonctions discontinues, et, si nous nous permettons, pour la facilité du calcul, d'y substituer une fonction analytique, il faudra que cette dernière soit choisie de telle sorte qu'elle tende rapidement vers 0 à partir de deux valeurs de Δ, l'une supérieure, l'autre inférieure à 0, et qu'en dehors de ces deux limites on puisse la considérer comme nulle.*

Il déduit de ceci deux propriétés essentielles de la fonction φ[19] :

> *Or la probabilité que l'erreur soit comprise entre Δ et une quantité Δ+dΔ qui en diffère infiniment peu sera exprimée par φ(Δ).dΔ et, par suite, la probabilité que l'erreur soit comprise entre D et D', par* $\int_{D}^{D'} \varphi(\Delta)\, d\Delta$.
>
> *Cette intégrale, prise depuis la plus grande valeur négative de Δ jusqu'à sa plus grande valeur positive, ou plus généralement depuis Δ = −∞ jusqu'à Δ = ∞, devra nécessairement être égale à 1. On aura donc* $\int_{-\infty}^{\infty} \varphi(\Delta)\, d\Delta = 1$.

---

[17] GAUSS, Carl Friedrich, *op. cit.*, pages 113-114.
[18] GAUSS, Carl Friedrich, *Ibid.*, page 114.
[19] GAUSS, Carl Friedrich, *Ibid.*, pages 114 -115.

Il revient à ses fonctions V, V', V", etc., d'inconnues p, q, r, s, etc., et détermine la valeur Ω qui correspond à la probabilité que toutes les valeurs M, M', M" résultent en même temps des observations. Il est à noter que Gauss considère que toutes les observations sont assimilées à des évènements indépendants les uns des autres[20].

> *Supposons donc qu'on ait un système déterminé de valeurs des quantités* p, q, r, s, *etc. : la probabilité que l'observation donnera pour V la valeur M sera exprimée par φ(M−V), après qu'on aura substitué dans V les valeurs de* p, q, r, s, *etc. ; de même φ(M '−V '), φ(M "−V "), etc. exprimeront les probabilités pour que les observations donnent aux fonctions V ', V ", etc., les valeurs M ', M ", etc. C'est pourquoi, tant qu'on pourra considérer toutes les observations comme des évènements indépendants les uns des autres, le produit*
>
> $$\varphi(M-V)\,\varphi(M'-V')\,\varphi(M''-V'')\ldots=\Omega$$
>
> *exprimera la probabilité que toutes ces valeurs résulteront en même temps des observations.*

Cette valeur de Ω et l'expression de φ vont conduire naturellement au résultat attendu. Gauss s'intéresse donc à la détermination de la fonction φ. Après une savante démonstration, il arrive au résultat suivant : $\varphi(\Delta)=x\,\mathrm{e}^{\frac{1}{2}k\Delta^2}$ où *x* et *k* désignent des constantes. Il poursuit alors ainsi[21] :

> *Or on voit facilement que la constante k doit être négative pour que Ω puisse devenir maximum : posons donc* $\frac{1}{2}k=-h^2$, *et comme d'après un élégant théorème de Laplace*[22], *on a* $\int_{-\infty}^{\infty}\mathrm{e}^{-h^2\Delta^2}\,d\Delta=\frac{\sqrt{\pi}}{h}$, *notre fonction deviendra* $\varphi(\Delta)=\frac{h}{\sqrt{\pi}}\mathrm{e}^{-h^2\Delta^2}$.

S'en suit un commentaire sur la validité de cette fonction[23] :

> *La fonction que nous venons de trouver ne peut pas exprimer, en toute rigueur, la probabilité des erreurs, puisque les erreurs possibles étant toujours renfermées entre certaines limites, la probabilité d'erreurs plus grandes devrait être toujours nulle, tandis que notre fonction a toujours une valeur finie. Cependant ce défaut, que présenterait également toute autre fonction analytique, n'a aucune importance dans les applications, parce que la valeur de notre fonction décroît si rapidement, pour peu que hΔ ait une valeur considérable, qu'on peut, en toute sûreté, la regarder comme équivalente à 0. D'ailleurs, la*

[20] GAUSS, Carl Friedrich, *Ibid.*, page 115.
[21] GAUSS, Carl Friedrich, *Ibid.*, page 119.
[22] Gauss et Laplace correspondaient régulièrement à propos de leurs travaux et avaient beaucoup d'estime l'un pour l'autre. On pourra se référer à l'ouvrage *De diversis artibus - Correspondance de Pierre Simon Laplace*, de Roger Hahn, publié chez Brepols.
[23] GAUSS, Carl Friedrich, *Ibid.*, page 119.

nature de la question ne permettra jamais d'assigner les limites des erreurs avec une rigueur absolue.

Le résultat tant attendu s'obtient alors de manière évidente. En effet, en posant $v = M - V$, $v' = M' - V'$, $v'' = M'' - V''$, etc... , la quantité Ω, définie ci-dessus, est égale à $\left(\frac{h}{\sqrt{\pi}}\right)^m \mathrm{e}^{-h^2\,(v^2+v'^2+v''^2+\ldots)}$ (m étant le nombre de fonctions V, V', V" considérées au départ) et elle est donc maximum lorsque la quantité $v^2+v'^2+v''^2+\ldots$ est minimale. Gauss énonce alors le principe et en souligne son importance[24] :

> *Donc le système de valeurs des inconnues* p, q, r, s*, etc., le plus probable correspond au cas où les carrés des différences entre les valeurs observées et les valeurs calculées des fonctions V, V', V", etc., donne la somme la plus petite possible, pourvu que toutes les observations soient également présumées précises.*
>
> *Ce principe, qui est de la plus grande utilité dans toutes les applications des mathématiques à la philosophie naturelle, doit être regardé comme un axiome, au même titre que le principe qui nous fait adopter la moyenne arithmétique des valeurs observées d'une même quantité comme la valeur la plus probable de cette quantité.*

LA POLÉMIQUE

Ce principe fondamental des statistiques, encore utilisé de nos jours, a fait l'objet d'une querelle de « paternité » entre ces deux grand savants. L'origine de celle-ci est le passage suivant du traité *Theoria motus corporum coelestium* de Gauss[25] :

Joseph BERTRAND (1822-1900)

> *Sed ex omnibus his principiis nostrum simplicissimum est, dum in reliquis ad calculos complicatissimos deferremur. Ceterum principum nostrum, quo iam inde ab anno 1795 usi summus, nuper etiam a clar. Legendre in opere* Nouvelles méthodes pour la détermination des orbites des comètes, Paris, 1806, *prolatum est, ubi plures aliae proprietates huius principii expositae sunt, quas hic brevitatis caussa supprimimus.*

Ce que Bertrand a traduit par[26] :

> *Mais de tous ces principes le nôtre est le plus simple, tous les autres nous entraînant dans des calculs extrêmement compliqués. Au reste, ce principe, dont nous avons fait usage dès l'année 1795, a été donné dernièrement par Legendre dans ses* Nouvelles méthodes pour la détermination des orbites des comètes, Paris, 1806 ; *on trouvera*

[24] GAUSS, Carl Friedrich, *Ibid.*, page 121.
[25] GAUSS, Carl Friedrich, *Theoria motus corporum coelestium*, 1809, page 121
[26] GAUSS, Carl Friedrich, *Méthode des moindres carrés*, traduit par J. Bertrand, 1855, page 133.

dans cet ouvrage plusieurs conséquences que le désir d'abréger nous a fait omettre.

Legendre n'a certainement pas apprécié que Gauss s'attribue la paternité de la méthode et il a été irrité par le fait que ce dernier dise qu'il l'utilisait depuis 1795, c'est-à-dire environ dix ans avant la parution de son propre ouvrage. Il envoya donc une lettre à Gauss le 31 mai 1809 dans laquelle il écrit ceci[27] :

> *Je ne vous dissimulerai donc pas, Monsieur, que j'ai éprouvé quelque regret de voir qu'en citant mon mémoire, page 221, vous dites* principum nostrum, quo iam inde ab anno 1795 usi sumus etc. *Il n'est d'aucune découverte qu'on ne puisse s'attribuer en disant qu'on avait trouvé la même chose quelques années auparavant ; mais si on n'en fournit pas la preuve en citant le lieu où on l'a publiée, cette assertion devient sans objet et n'est plus qu'une chose désobligeante pour le véritable auteur de la découverte. En Mathématiques il arrive très souvent qu'on trouve les mêmes choses qui ont été trouvées par d'autres et qui sont bien connues ; c'est ce qui m'est arrivé nombre de fois, mais je n'en ai point fait mention et je n'ai jamais appelé* principum nostrum *un principe qu'un autre avait publié avant moi. Vous êtes assez riche de votre fonds, Monsieur, pour n'avoir rien à envier à personne ; et je suis bien persuadé au reste que j'ai à me plaindre de l'expression seulement et nullement de l'intention…*

Il semble qu'il n'y ait pas eu de réponse de la part de Gauss. Cependant, dans un courrier[28] daté du 25 novembre 1811, Laplace demande des explications à Gauss à ce sujet :

> *Monsieur Gauss dit dans son ouvrage sur le mouvement elliptique qu'il la possédait avant que M. Legendre l'ait publiée ; je désirerais bien savoir si avant cette publication on avait imprimé quelque chose en Allemagne sur cette méthode et je prie Monsieur Gauss de bien vouloir m'en instruire.*

Gauss lui répond dans une lettre datée du 30 janvier 1812, dans laquelle il revendique, en quelque sorte, la paternité de la méthode et s'étonne de la réaction de Legendre[29].

> *J'ai fait usage de la méthode des moindres carrés depuis l'an 1795 et je trouve dans mes papiers que le mois de Juin 1798 est l'époque où je l'ai rapprochée aux principes du calcul des probabilités : une note la dessus se trouve dans un journal que j'ai tenu sur mes occupations mathématiques depuis l'an 1796, et que j'ai montré dans ces jours à Mr de Lindenau.*
>
> *Cependant mes applications fréquentes de cette méthode ne datent que de l'année 1802, depuis ce temps j'en fais usage pour ainsi dire tous les jours dans mes calculs astronomiques sur les nouvelles planètes. Comme je m'étais proposé depuis ce temps de réunir toutes les méthodes dont je me suis servi dans un ouvrage étendu[30], (que j'ai commencé en 1805 et dont le Manuscrit d'abord en*

---

[27] GAUSS, Carl Friedrich, *Werke , Band 10/1*, page 380. Ce document est disponible sur le site suivant : ***gdz**.sub.uni-goettingen.de/* (site de : Göttinger Digitalisierungszentrum)
[28] GAUSS, Carl Friedrich, *Ibid.*, page 380.
[29] GAUSS, Carl Friedrich, *Ibid.*, pages 373-374.
[30] Il s'agit de son ouvrage *Theoria motus corporum coelestium.*

allemand était achevé en 1806, mais lequel à la prière de Mr. Perthes j'ai traduit depuis en latin ; l'impression a commencé en 1807 et n'est finie qu'en 1809), je ne me suis pas hâté d'en publier un morceau détaché, ainsi Mr. Legendre m'est prévenu. Au reste j'avais déjà communiqué cette même méthode, beaucoup avant la publication de l'ouvrage de M. Legendre, à plusieurs personnes, en autres à Mr. Olbers en 1803 qui certainement doit se le rappeler. Ainsi, pouvais je dans ma théorie des mouvements des planètes, *parler de la méthode des moindres quarrés, dont j'avais fait depuis 7 ans mille et mille applications, dont j'avais développé la théorie, dans la section 3^me^ du II.^livre^ de cet ouvrage, du moins en allemand, beaucoup avant d'avoir vu l'ouvrage de Mr. Legendre - je dis, pouvais je parler de ce principe, que j'avais annoncé à plusieurs de mes amis déjà en 1803 comme devant faire partie de l'ouvrage que je préparais - comme d'une méthode empruntée de Mr. Legendre ? Je n'avais pas l'idée, que Mr. Legendre pouvait attacher tant de prix à une idée aussi simple, qu'on doit plutôt s'étonner qu'on ne l'a pas eue depuis 100 ans, pour se fâcher que je raconte, que je m'en suis servi avant lui ? En effet il serait très facile de le prouver à tout le monde par des témoignages qu'on ne saurait refuser, si cela valait la peine. Mais j'ai cru que tous ceux qui me connaissent le croiraient même sur ma parole, ainsi que je l'aurais cru de tout mon cœur si Mr. Legendre avait avancé, qu'il avait possédé la méthode déjà avant 1795. J'ai dans mes papiers beaucoup de choses dont peut être je pourrai perdre la priorité de la publication : mais soit, j'aime mieux faire mûrir les choses.*

Ces lettres ne permettent pas de désigner le « père » de la méthode, mais est-ce bien important ? Ils sont tous deux de grands savants reconnus. Il est cependant certain que Legendre a été le premier à la publier, comme le reconnaît Gauss.

APPLICATION DE LA MÉTHODE

Suite à son exposé sur la méthode des moindres carrés, Legendre en propose une application à la mesure des degrés du méridien.

Il commence par des préliminaires permettant d'obtenir les équations aux erreurs[31].

Supposant que le méridien est une ellipse dont les axes sont dans le rapport de 1 à $1+\alpha$, si on désigne par D la longueur du 45^ème^ degré, et par S celle de l'arc compris entre les deux latitudes L et L', on aura par les formules connues, et en exprimant $L'-L$ en degrés :

$$\mathrm{S}=\mathrm{D}(\mathrm{L}'-\mathrm{L})-\frac{3}{2}\alpha\mathrm{D}.\frac{180}{\pi}\sin(\mathrm{L}'-\mathrm{L})\cos(\mathrm{L}'+\mathrm{L})\ ;$$

d'où résulte : $$\mathrm{L}'-\mathrm{L}=\frac{\mathrm{S}}{\mathrm{D}}+\frac{3}{2}\alpha.\frac{180}{\pi}\sin(\mathrm{L}'-\mathrm{L})\cos(\mathrm{L}'+\mathrm{L}).$$

[31] LEGENDRE, Adrien-Marie, *op. cit.*, 1805, page 76.

Comme le 45ème degré est d'environ 28500 modules, égaux chacun à deux toises, on peut faire $\dfrac{1}{\mathrm{D}} = \dfrac{1+\varepsilon}{28500}$, ε *étant une fraction très petite*[32]*, et on aura*

$$\mathrm{L'} - \mathrm{L} = \frac{\mathrm{S}}{28500} + \varepsilon.\frac{\mathrm{S}}{28500} + \alpha.\frac{270}{\pi}\ \sin(\mathrm{L'}-\mathrm{L})\ \cos(\mathrm{L'}+\mathrm{L}), \qquad (a)$$

équation qui pour chaque arc dont on connaît la longueur avec la latitude de ses extrémités, donnera une relation entre α et ε.

Il donne ensuite un tableau[33] comportant les longueurs de différents arcs de la méridienne de France ainsi que les latitudes des parallèles qui les séparent telles que, dit-il, « elles résultent de l'opération exécutée par les célèbres astronomes Delambre et Méchain. »

Lieu de l'observation.	*Sa latitude.*	*Arcs compris exprimés en modules.*	L' — L	L' + L
Dunkerque	51° 2' 10" 50			
		DP 62472.59	2° 11' 20" 75	99° 53' 0"
Panthéon à Paris	48. 50 49. 75			
		PE 76145. 74	2 40 7. 25	95 1 32
Evaux..............	46 10 42. 50			
		EC 84424. 55	2 57 48. 10	89 23 37
Carcassonne.....	43 12 54. 40			
		CM 52749. 48	1 51 9. 60	84 34 39
Montjouy........	41 21 44 80			

Dans ce tableau, la 3ème colonne correspond aux différentes valeurs de S.

À partir des données ci-dessus, on obtient le tableau suivant, les résultats étant arrondis à 10^{-3}.

S	$\dfrac{\mathrm{S}}{28500}$	$\dfrac{270}{\pi}\ \sin(\mathrm{L'}-\mathrm{L})\ \cos(\mathrm{L'}+\mathrm{L})$
DP	2,192	−0,563
PE	2,672	−0,351
EC	2,962	0,047
CM	1,851	0,263

En substituant successivement ces valeurs dans l'équation (a), on obtient quatre équations d'inconnues α et ε. Legendre fait alors le commentaire suivant[34] :

---

[32] Dans son ouvrage, Legendre utilise le caractère pour désigner cette fraction. Ne l'ayant pas trouvé dans les polices, je l'ai remplacé dans le texte par epsilon (ε).
[33] LEGENDRE, Adrien-Marie, *Ibid.*, page 76.

Mais comme ces quatre équations ne peuvent être satisfaites toutes à la fois, nous supposerons qu'elles ont lieu en attribuant une certaine erreur à la latitude de chaque lieu, et nous appellerons E', E", etc les corrections additives aux latitudes de Dunkerque, du Panthéon, etc. Ces erreurs n'entrent que dans le premier membre de chaque équation : elles sont trop petites pour affecter le terme multiplié par α *dans le second membre. Voici donc les équations qui résultent des quatre arcs mesurés dans l'opération de la méridienne*

$$\begin{aligned}
E' - E'' &= 0.002923 + \varepsilon(2.192) - \alpha(0.563) \\
E'' - E''' &= 0.003100 + \varepsilon(2.672) - \alpha(0.351) \\
E''' - E^{\text{IV}} &= -0.001096 + \varepsilon(2.962) + \alpha(0.047) \\
E^{\text{IV}} - E^{\text{V}} &= -0.001808 + \varepsilon(1.851) + \alpha(0.263)
\end{aligned}$$

Legendre n'indique pas dans son ouvrage comment il a obtenu les valeurs 0.002923, 0.003100, –0.001096 et –0.001808 figurant dans les équations.

Ensuite, il fixe E''' comme inconnue, apparemment de façon arbitraire, et obtient alors les cinq équations suivantes[35] :

$$\begin{aligned}
E' &= E''' + 0.006023 + \varepsilon(4.864) - \alpha(0.914) \\
E'' &= E''' + 0.003100 + \varepsilon(2.672) - \alpha(0.351) \\
E''' &= E''' \\
E^{\text{IV}} &= E''' + 0.001096 - \varepsilon(2.962) - \alpha(0.047) \\
E^{\text{V}} &= E''' + 0.002904 - \varepsilon(4.813) - \alpha(0.310)
\end{aligned}$$

Il écrit que la somme des erreurs doit être nulle, ce qui donne l'équation ci-dessous :

$$0 = 5\,E''' + 0{,}013123 - 0{,}239\,\varepsilon - 1{,}622\,\alpha$$

c'est-à-dire : $E''' = -0{,}002625 + 0{,}048\,\varepsilon + 0{,}324\,\alpha$.

Il substitue cette valeur dans les équations précédentes et obtient le système suivant[36] :

$$\begin{aligned}
E' &= 0.003398 + \varepsilon(4.912) - \alpha(0.590) \\
E'' &= 0.000475 + \varepsilon(2.720) - \alpha(0.027) \\
E''' &= -0.002625 + \varepsilon(0.048) + \alpha(0.324) \\
E^{\text{IV}} &= -0.001529 - \varepsilon(2.914) + \alpha(0.277) \\
E^{\text{V}} &= 0.000279 - \varepsilon(4.765) + \alpha(0.014)
\end{aligned}$$

Il poursuit en rappelant le principe de sa méthode[37]:

---

[34] LEGENDRE, Adrien-Marie, *Ibid.*, page 77.
[35] LEGENDRE, Adrien-Marie, *Ibid.*, page 77.
[36] LEGENDRE, Adrien-Marie, *Ibid.*, page 78.

> *Pour exprimer ensuite la condition du minimum par rapport à ε, il faut multiplier la première équation par 4,912 coefficient de ε, la seconde par 2,720, la troisième par 0,048, la quatrième par –2,914, la cinquième par –4,765 et égaler à 0 la somme de tous les produits. On opèrera semblablement par rapport à α.*

Il obtient ainsi un système de deux équations à deux inconnues α et ε [38]:

$$\begin{aligned} 0 &= 0.020983 + \varepsilon(62.726) - \alpha(3.830) \\ 0 &= -0.003287 - \varepsilon(3.830) + \alpha(0.531) \end{aligned}$$

La résolution de ce système conduit à $\alpha = 0{,}00675$ et $\varepsilon = 0{,}0000778$.

L'aplatissement α est donc égal à $\frac{1}{148}$ et le 45ème degré vaut alors $\frac{28500}{1+\varepsilon}$ c'est-à-dire 28497,78.

Legendre termine en commentant ces valeurs qui ne sont pas conformes aux valeurs connues, ces variations étant dues à des anomalies de latitude qui, selon lui, ne doivent pas être attribuées aux observations mais à des attractions locales qui agissent sur le fil à plomb et aux inégalités observées dans les azimuts.

CONCLUSION

Quelle qu'en soit la paternité, cette méthode, d'une grande simplicité comme l'ont fait remarquer Gauss et Legendre, reste une des plus grandes avancées réalisées dans le domaine des statistiques. Elle est toujours d'actualité, même si, comme il est dit dans l'introduction, elle tend à être supplantée par celle du maximum de vraisemblance.[39]

BIBLIOGRAPHIE

GAUSS, Carl Friedrich, *Méthode des moindres carrés. Mémoires sur la combinaison des observations*, traduit par Joseph BERTRAND. Paris, Mallet-Bachelier, 1855.

GAUSS, Carl Friedrich, *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Hamburg, Perthes et Besser, 1809.

GAUSS, Carl Friedrich, *Werke*, *Zehnten Bandes Erste Abteilung*. Göttingen, Königlichen Gesellschaft der Wissenschaften, 1917.

LEGENDRE, Adrien-Marie, *Nouvelles méthodes pour la détermination des orbites des comètes*. Paris, Firmin Didot, 1805.

QUÉTELET, Adolphe, *Recherches statistiques*. Bruxelles, Hayez, 1844

---

[37] LEGENDRE, Adrien-Marie, *Ibid.*, page 78.
[38] LEGENDRE, Adrien-Marie, *Ibid.*, page 78.
[39] Pour des compléments sur cette méthode, on pourra consulter l'œuvre de Ronald Fisher, « On the mathematical foundations of theoretical statistics », *Philosophical Transactions of the Royal Society*, n° 222, 1922, p. 309-368.