

LA COURBE DE GAUSS OU LE THÉORÈME DE BERNOULLI RACONTÉ AUX ENFANTS

Bernard BRU¹

RÉSUMÉ – *Ce texte est à visée essentiellement pédagogique. Il présente les principales étapes de la construction de la « loi des erreurs » de Gauss, de la fin du XVII^e siècle (Jacques Bernoulli) au tout début du XX^e siècle.*

Il met en évidence l'articulation entre les idées successives qui ont abouti à la conception actuelle de cette loi, telle qu'elle est enseignée aux élèves s'initiant au calcul des probabilités.

MOTS-CLÉS – J. Bernoulli, A. de Moivre, P.S. Laplace, C.F. Gauss, A. Quetelet, J. Bertrand, Courbe de Gauss, Loi binomiale, Formule de Stirling

SUMMARY – The Gaussian curve or the Bernoulli theorem told to children
The purpose of this paper is mainly educational. It is an introduction to the main stages of the elaboration of the Gaussian "Law of error", from the end of the XVIIth Century (Jacob Bernoulli) to the beginning of the XXth Century.

It shows the logical links between the successive ideas which would result in the actual conception of this law, such as presented in the current elementary teaching of probabilities.

KEY-WORDS – J. Bernoulli, A. de Moivre, P.S. Laplace, C.F. Gauss, A. Quetelet, J. Bertrand, Gaussian curve, Binomial law, Stirling's formula

Avertissement de la rédaction

La rédaction ne disposant pas de dessinateur assez habile pour réaliser les figures et dessins suggérés par B. Bru, nous les remplaçons par les indications écrites par l'auteur lui-même. Chaque lecteur pourra ainsi se représenter, ou, s'il a assez de talent, dessiner selon son imagination et à sa fantaisie, chacune de ces figures.

Le « texte qui suit a été écrit à la demande des organisateurs de l'exposition « Pourquoi les maths ? », sous l'égide de l'Unesco (<http://www.MathEx.org>). Il avait pour but de donner quelques idées de présentation du théorème de Bernoulli et de la courbe de Gauss. Je remercie très chaleureusement Marc Barbut, éditeur de la Revue *Mathématiques et Sciences humaines* de m'avoir encouragé à le publier séparément dans sa revue et Mireille Chaleyat-Maurel, responsable de l'exposition, de m'avoir si gentiment autorisé à le faire.

¹ MAP5, Université Paris V, 45 rue des Saints-Pères 75006 Paris, bernard.bru@univ-paris5.fr

L'auteur aimerait dédier ce petit texte à Christelle Balloy, professeur de mathématiques au Collège Boris-Vian de Lille, en témoignage d'amitié.

D'abord, Gauss aurait été assez mécontent d'apprendre qu'on l'affublerait de cette courbe, et qu'il serait principalement connu dans le public pour cette raison. Pourquoi pas le fil à couper le beurre ? Mais passons.

Dans le cas le plus simple, l'équation de la courbe de Gauss s'écrit :

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Quelque chose de compliqué, qui contient les deux nombres les plus mystérieux des mathématiques, $\pi = 3,14159\dots$, le rapport de la circonférence au diamètre, et $e = 2,71828\dots$, la base des logarithmes népériens, des nombres qu'on trouve sur toutes les calculettes mais dont l'écriture décimale ne s'arrête jamais et ne contient pas de périodicité. Voilà qui commence mal.

La forme de la courbe de Gauss en revanche est très connue. Tout le monde l'a vue un jour ou l'autre dans un journal. C'est une cloche plate ou un chapeau de gendarme vu de face, lorsque les gendarmes en portaient un de cette forme-là.

Figure 1. La courbe de Gauss standard, en gros, en marquant bien les unités sur l'axe des x (éventuellement, un boa qui a avalé un dromadaire)

Une propriété importante de cette courbe, outre sa symétrie visible, est que la surface qu'elle délimite avec l'axe horizontal est égale à 1. Un résultat qui n'a rien d'évident et qui remonte à Laplace et Euler vers 1777 (beau cadeau de baptême pour le petit Gauss qui vient de naître à Brunswick). Il s'agit donc de ce qu'on appelle une densité de probabilité, ce qui s'écrit :

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

On peut déplacer le point le plus haut de la courbe, en changeant la variable x en $x-a$. l'axe de symétrie de la courbe se trouve alors à l'abscisse a . On peut aussi évaser ou rétrécir la cloche en multipliant l'exposant de l'exponentielle par un paramètre positif convenable. Pour nous conformer à l'usage, notons le $\frac{1}{\sigma^2}$. Plus σ (sigma) est grand plus la courbe est évasée, plus il est petit plus elle est mince. En faisant les deux opérations en même temps, on peut fabriquer toute une famille de courbes de Gauss, d'équations :

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

qui sont autant de cloches de la forme et à l'emplacement qu'on voudra et qui possèdent toutes la même propriété d'être des densités de probabilité : la surface qu'elles délimitent avec l'axe des x est égale à un. On peut remarquer (si on sait dériver deux fois) que les points d'inflexion de la cloche, là où elle change sa façon de voir le monde, sont situés à une distance de l'axe de symétrie égale à σ . On observe enfin qu'au-delà du double de σ , la courbe est tellement collée à l'axe des x qu'on ne la voit plus du tout. Plus de 95 % de la surface de la cloche se trouve comprise entre plus et moins deux sigmas de l'axe de symétrie. Donc tout ou presque se passe dans cet intervalle-là. Pour cette raison σ se nomme l'écart-type de la courbe de Gauss. C'est avec a , l'un des deux paramètres caractéristiques de la courbe.

Figure 2. Des courbes de Gauss de formes variées, par exemple une d'écart-type $\frac{1}{2}$ et une autre d'écart-type 2 avec des origines différentes (éventuellement des boas ou des cloches ou des chapeaux de gendarme de formes variées)

Voilà, nous savons tout sur la courbe de Gauss.

Et en même temps nous ne savons rien. Parce qu'une courbe en chapeau de gendarme n'a pas besoin de tant de sophistication mathématique pour être tracée. On en dessine de très belles sans π ni e ni rien de tout ça. Un boa qui aurait avalé de travers une cloche ou un dromadaire (pas un chameau) ferait aussi bien l'affaire. Pourtant on nous dit que la courbe de Gauss est un des piliers de la statistique. On la trouve dans tous les carnets de santé et dans mille autre choses encore, irruption singulière de l'analyse transcendante dans le panier de la ménagère.

Bizarre !

Voyons cela de plus près.

D'abord le point de vue du mathématicien. Parce que c'est certainement un mathématicien qui est à l'origine de cette fonction biscornue. Un ou plusieurs, et pourquoi pas Gauss ?

Donc première question : quelle est l'origine mathématique de la courbe de Gauss ?

Réponse : dans le calcul des chances, au jeu de pile ou face vers 1730 (la mère du petit Carl Friedrich n'était pas encore née).

On joue, mettons, cent fois à pile ou face. Tout le monde sait qu'on va obtenir sensiblement le même nombre de piles que de faces, parfois un peu plus de piles que de faces, parfois un peu moins, mais en tout cas pas de très grands écarts, sauf exceptions rares. Peut-on préciser davantage cet énoncé de sens commun ? Quels écarts sont raisonnables ? Quelles sont leurs degrés de rareté, leurs « probabilités » ? C'est justement l'objet du calcul des chances ou des probabilités de répondre à ce genre de questions dans un cadre mathématique.

Le premier auteur à s'être posé la question à peu près dans ces termes paraît être Jacques Bernoulli, à Bâle, vers 1677 (les grands-mères de Gauss n'étaient pas encore nées, pour les arrière-grands-pères, on ne sait pas).

Considérons par la pensée toutes les suites possibles de cent parties de pile ou face, par exemple cent piles à la suite, face d'abord et 99 piles ensuite, pile d'abord, face ensuite et 98 piles, etc., il y a 2^{100} suites de cette sorte, un nombre gigantesque : 1 suivi de 30 chiffres entre 0 et 9).

Figure 3. Quelques suites de parties de pile ou face imagées

Considérons maintenant les suites contenant un nombre donné N de piles. N peut prendre toutes les valeurs entières entre 0 à 100, mais on imagine bien qu'il n'y aura pas le même nombre de suites pour toutes ces valeurs. Par exemple, pour $N = 0$, il n'y en a qu'une, la pièce doit retomber sur face à tous les coups, on peut dire (c'est une définition) que sa chance de se produire est 1. Pour $N = 1$, il y en a cent, parce que pile peut se trouver en cent places différentes : la chance d'obtenir exactement une fois pile est 100. Pour $N = 2$, c'est un peu plus compliqué, il y a deux places possibles pour pile parmi les cent jets de la pièce ; c'est un problème combinatoire déjà traité par les Arabes, les Chinois, les Indiens, ..., bien avant les savants de la Renaissance. En tout cas, le petit Gauss, au berceau, aurait répondu immédiatement, si sa maman lui avait posé la question : il y en a 4950. Ce qui paraît beaucoup mais est, en réalité, absolument invisible parmi les 2^{100} suites de parties possibles. Ensuite les choses se compliquent davantage, non pas que la formule combinatoire soit difficile à établir, mais parce qu'elle ne s'évalue pas facilement. Par exemple pour $N = 50$, le nombre de suites possibles, où figurent 50 fois pile et 50 fois face, est égal au nombre de façons de choisir 50 choses parmi 100, c'est-à-dire :

$$\frac{1.2.3.....99.100}{(1.2.3.....49.50)^2} = \frac{100}{50} \cdot \frac{99}{49} \cdot \frac{98}{48} \cdot \frac{52}{2} \cdot \frac{51}{1} \ggg 2^{50}$$

C'est beaucoup. Quoique certainement plus petit que 2^{100} , le nombre total des suites. De combien est-ce plus petit ? la moitié, le tiers, ... ? Difficile à dire. Pourtant si on pouvait dire quelque chose d'assuré à ce sujet, on répondrait à la question de Bernoulli. On affecterait à chaque valeur de N un nombre, le nombre de suites qui lui correspondent, son nombre de « chances », dont l'importance nous préciserait en un certain sens, mathématique, la possibilité qu'elle a de se produire. Et l'on prédirait, on conjecturerait, ce qui va se passer à l'aide de sa chance calculée. Quitte à laisser ensuite à l'expérimentateur, au praticien et au banquier le soin de vérifier si, dans cette affaire, le mathématicien déraisonne complètement ou s'il a vu juste.

Quoi qu'il en soit, les nombres de chances de N vont visiblement en croissant de $N = 0$, jusqu'à $N = 50$, et ensuite décroissent symétriquement, jusqu'à valoir 4950 pour $N = 98$, 100 pour $N = 99$, et 1 pour $N = 100$. Si on représente ces nombres par des « petits » bâtons verticaux d'abscisse 0, 1, 2, ..., 100, que voit-on ? On doit obtenir une sorte d'arête de poisson, comme une demi-sole dont on aurait mangé la chair, un squelette de cloche en somme, une courbe de Gauss squelettique. Si on n'aime pas les bâtons minces, on peut imaginer qu'ils ont une épaisseur unité et qu'ils se touchent, on obtient ainsi une cloche en escaliers qui montent, puis qui descendent.

Figure 4. Une cloche en bâtons et une cloche en escaliers, si possible personnalisées

Essayons de tracer plus précisément cette courbe en bâtons, ou en escalier, sur une feuille de papier. La question de l'échelle se pose. On passe de 1 à 100 puis à 4950 et après on ne sait pas trop mais on atteint très vite les milliards. Pour faire tenir le dessin dans la page, il faut que l'échelle verticale soit considérablement rétrécie. Le plus simple est de diviser chacun des nombres de chances par leur somme totale 2^{100} , c'est-à-dire faire un diagramme de probabilités : la somme des hauteurs des bâtons est égale à un ; ou encore si on suppose que les bâtons ont une épaisseur unité, et qu'ils se touchent, la surface totale des bâtons est égale à 1. Mais alors les premiers bâtons ne se voient plus du tout. Ils se confondent avec l'axe horizontal. Quand vont-ils commencer à se voir ? Jusqu'où vont-ils monter ? On est revenu au problème de l'évaluation numérique des chances, rapportées au nombre total 2^{100} . Par exemple, pour $N = 50$, que vaut la fraction :

$$\frac{100.99.....52.51}{2^{100}.50.49.....2.1}$$

Même question pour $N = 49$, $N = 48$, ...

Jacques Bernoulli n'est pas parvenu à résoudre complètement ce problème, et l'on peut être assuré qu'il a cherché longtemps. En revanche, à l'aide de majorations ingénieuses (et d'ailleurs suffisantes pour les applications), il est arrivé à démontrer que si on additionne tous les nombres de suites, toutes les chances, correspondant à des N assez voisins de 50, on obtient une écrasante majorité de l'ensemble de toutes les parties possibles. Les parties à peu près équitables ont infiniment plus de chances que les autres. En fait, Jacques Bernoulli montre que si l'on considère non plus 100 parties, mais un très grand nombre n de parties, si on note toujours N le nombre de piles obtenus au cours de ces n parties, pour n assez grand, on peut parier un million d'euros contre 1 que l'écart entre la fréquence des piles, N/n , et $1/2$ ne dépassera pas $1/100$, par exemple. Cet énoncé figure dans un ouvrage de J. Bernoulli publié en 1713, après sa mort survenue en 1705, et intitulé *Ars conjectandi*, l'Art de conjecturer, tout un programme ! Jacques Bernoulli appelait son résultat le théorème d'or et il en était infiniment fier. C'est le premier énoncé mathématique de ce qui s'appellera à partir de 1835, « la loi des grands nombres ». Toutefois, on l'a dit, Jacques Bernoulli n'est pas parvenu à préciser tout à fait son résultat. Il aurait fallu pour cela évaluer par exemple la fraction écrite ci-dessus, dont le numérateur et le dénominateur sont des nombres qui défient l'imagination, et toute autre fraction semblable où l'on voit apparaître, en haut et en bas, les produits successifs des n premiers entiers, pour un entier n trop grand pour un calcul direct. Et ce « trop grand » n'est pas si grand. Déjà pour $n = 13$, (si la calculatrice de mon ordinateur marche bien) on dépasse les six milliards :

$$1.2.3.4.5.6.7.8.9.10.11.12.13 = 6.227.020.800$$

Il faut donc une formule approchée qui donne à fort peu près le produit cherché et permette de calculer la ou les fractions de Bernoulli. Cette formule se trouve maintenant dans tous les cours de mathématiques. C'est la formule de Stirling qui a été inventée par Abraham de Moivre et James Stirling, notamment dans le dessein d'évaluer les chances au jeu de pile ou face. Moivre est un huguenot français, né à Vitry-le-François en 1667. Après la révocation de l'Édit de Nantes, il se réfugie à Londres où il s'établit comme mathématicien, l'un des plus grands de son temps. C'est le mathématicien qui a le plus contribué au calcul des chances entre 1710 et 1756, date de sa mort. La formule de Stirling-Moivre de 1730 s'écrit :

$$1.2.3.....n \cong \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n$$

Où l'on voit apparaître π et e .

Par exemple pour $n = 13$, le terme de droite de l'égalité approchée précédente vaut à peu près :

$$6.187.446.400$$

Remarque. Je n'ai que la calculette de l'ordinateur c'est-à-dire les 4 opérations. J'ai donc fait des arrondis très arrondis, et je ne garantis pas la justesse des 9 derniers chiffres, qu'il serait bon de vérifier.

Ce qui est très bon en valeur relative, la seule importante pour le calcul que nous voulons faire : le quotient de la vraie valeur par la valeur approchée est égal à 1,006. Et plus n augmente, plus la formule devient bonne de cette façon : le quotient des deux termes s'approche de 1 de plus en plus (mais pas la différence qui, elle, augmente indéfiniment avec n).

Magnifique formule donc, qui sera encore précisée par Euler au milieu du XVIII^e siècle, puis redémontrée de tout autre façon vers 1780 par Laplace, qui y verra un cas très particulier d'une formule générale permettant d'évaluer les intégrales définies des fonctions contenant des facteurs élevés à de grandes puissances, parce que le produit des n premiers entiers peut en effet se voir comme une intégrale définie de cette sorte, la fonction gamma d'Euler.

$$1.2.3.....n = \int_0^{\infty} e^{-x} x^n dx$$

Pour comprendre la formule de Stirling, il suffit de dessiner la courbe $e^{-x} x^n$ entre 0 et l'infini. Dès que n est grand, elle ressemble à une cloche très mince centrée à l'abscisse n , où elle vaut précisément $\left(\frac{n}{e}\right)^n$, d'où il résulte que la valeur du produit des n premiers entiers est égale à fort peu près à la surface délimitée par cette cloche, c'est-à-dire à sa hauteur multipliée par sa largeur, qui est nécessairement petite par rapport à n . On se ramène facilement, à l'aide d'un changement de variable au voisinage de n , à la courbe en cloche véritable, où l'on retrouve π et la racine carrée de n . Voilà la démonstration de Laplace, qui présente cet avantage de pouvoir s'étendre assez pour s'adapter au théorème général de Laplace, dont nous parlons ci-dessous.

Mais nous nous égarons.

Muni de la formule de Stirling, il est alors facile d'évaluer la fraction exprimant les probabilités au jeu de pile ou face prolongé aussi longtemps qu'on veut. C'est précisément ce qu'a fait Moivre. En particulier, la probabilité d'obtenir exactement 50 piles, en 100 jets de la pièce, devient, avec cette approximation :

$$\frac{100 \cdot 99 \cdot \dots \cdot 52 \cdot 51}{2^{100} \cdot 50 \cdot 49 \cdot \dots \cdot 2 \cdot 1} \cong \frac{1}{\sqrt{50\pi}} \cong 0,08$$

Soit 8 % des cas (et non pas la moitié ou le tiers conjecturés plus haut).

On peut, de la même façon, calculer les autres probabilités. Ce qui est un peu long, d'autant que, disposant de la courbe de Gauss, on peut résumer tous les calculs en remarquant que la formule de Stirling, convenablement exploitée, permet d'approcher de près les bâtons ou l'escalier inconnus des 100 jets de la pièce par une courbe de Gauss dont les éléments caractéristiques sont :

$$a = 50, \quad \sigma = 5$$

L'escalier inconnu, vu d'assez loin pour éliminer les détails des marches est la courbe de Gauss correspondant à un écart-type de 5 et un axe de symétrie d'abscisse 50, le nombre de piles qui a le plus de chances pour lui. D'où il résulte, d'après ce que nous savons déjà sur la courbe de Gauss, que les chances d'obtenir moins de 40 piles ou plus de 60 sont faibles (moins de 5 %) et que tout se passe entre 40 et 60, l'escalier s'infléchissant aux abscisses 45 et 55 et culminant à l'abscisse 50 où il atteint la hauteur 8 %.

Figure 5. Une cloche centrée en 50, d'écart-type 5, qui monte à 8

Ce n'est pas vraiment une cloche, plutôt un phare ou un boa qui aurait avalé une horloge comtoise dans le mauvais sens. Quoi qu'il en soit, tout est dit pour le cas de 100 parties.

Le cas de n parties se traite évidemment de la même façon. On obtient ainsi le « théorème de Moivre » : si on joue un grand nombre n de fois à pile ou face, les chances des différents nombres de piles obtenues se distribuent à fort peu près comme une courbe de Gauss dont l'écart-type est $\frac{\sqrt{n}}{2}$ et dont l'axe de symétrie se trouve à l'abscisse $n/2$.

Par exemple si $n = 1.000.000$, l'écart-type est égal à 500. La partie de la courbe qui se détache de l'axe horizontal a une largeur inférieure à 2000, c'est-à-dire la cinq centième partie de la largeur totale de 1.000.000. Le dessin, à l'échelle d'une feuille de papier ordinaire, ressemble à un manche à balai très mince (moins d'un demi-millimètre de diamètre) situé au centre de la feuille. Le théorème de Bernoulli prend maintenant une forme numérique simple qui se prête facilement aux applications, si on dispose d'une table de la courbe de Gauss.

Indubitablement la courbe de Gauss est sortie des limbes où elle se tenait cachée, à Londres, vers 1730, et c'est vraisemblablement Abraham de Moivre qui l'a aperçue le premier, sans toutefois la tracer ni se rendre compte qu'elle définissait une sorte de distribution de probabilités idéale.

Figure 6. On pourrait dessiner Big Ben marquant 1730 et une courbe en cloche en forme de fantôme qui volette autour sous l'œil étonné d'un personnage en perruque marquée Moivre.
Mais cela risque d'être compliqué. Ou encore une courbe en cloche en forme de bateau à voile voguant sur la Tamise, avec le capitaine Moivre à la barre, etc.

Le théorème de Moivre, comme celui de Bernoulli, s'étend sans grande difficulté au cas des pièces dissymétriques, qui ont tendance à favoriser un côté plutôt que l'autre. De sorte qu'il s'applique couramment à un grand nombre de situations concrètes, par exemple au calcul des « fourchettes » des instituts de sondages et du second tour des élections présidentielles. Mais n'anticipons pas.

Nous sommes encore loin du panier de la ménagère moderne.

Nous en sommes restés à l'année 1730. Nous allons sauter quelques étapes pour ne pas être trop long. Le calcul des chances s'est d'ailleurs quelque peu assoupi au XVIII^e siècle et Moivre n'a guère été suivi, ni lu. Toutefois on assiste à un sursaut remarquable de la théorie au cours des années 1770 (L'âme rationnelle du petit Carl Friedrich aurait-elle estimée qu'il était grand temps de naître ?). Il ne s'agit alors plus de calculer les chances des jeux de hasard et de fixer les enjeux des paris et les primes d'assurance, mais de permettre à l'Astronomie de précision, le triomphe du XVIII^e siècle savant, d'atteindre à une précision numérique jamais égalée, qui permette de concilier les résultats des observations avec la Mécanique céleste censée en rendre compte. On ne peut imaginer d'ambition intellectuelle plus noble et les plus grands savants du temps vont rivaliser de génie sur un sujet mathématique nouveau, la théorie mathématique des erreurs d'observation. Une théorie probabilisée : lorsqu'on fait une mesure à l'aide d'un instrument optique, même si on s'affranchit des biais systématiques de l'appareil, il reste les erreurs inévitables, défauts d'attention, variations anarchiques de la réfraction atmosphérique, vibrations diverses, etc. autant de causes accidentelles d'erreurs qui affecteront la précision de la mesure. Si on recommence la même mesure dix fois, cent fois, etc., on n'obtient jamais la même valeur. Et que dire du cas où les observations sont indirectes et ne donnent le résultat cherché (par exemple la masse de Jupiter) que par l'intermédiaire d'équations mettant en jeu plusieurs mesures de natures variées.

Restons dans le cas le plus simple, n mesures effectuées sous les conditions les plus égales possible. Cela ressemble fort à un jeu de hasard dont les résultats varient d'une partie à l'autre. Ce n'est plus une série de piles et de faces mélangées, mais une série de nombres à peu près égaux, sans plus, tantôt plus grands, tantôt plus petits, sans choix visibles. Imaginons que n soit très grand. Et représentons sur un dessin ces mesures à la manière de ce qui a été fait plus haut. Sur un axe horizontal, on trace des subdivisions égales à l'unité et au-dessus de chacune d'entre elles on dessine un rectangle de hauteur égale au nombre de mesures faites qui tombent dans l'intervalle en question. Plutôt que les nombres absolus il est bien préférable de mettre les nombres relatifs, rapportés à n . On obtient alors un dessin en forme d'escalier qui monte et qui descend un peu n'importe comment et dont la surface qu'il délimite est égale à 1. Vu de loin, cela ressemble à une courbe avec des creux et des bosses, dont la forme, si on suppose n immense, est d'une certaine façon caractéristique de notre procédé de mesure. Si l'instrument est très précis, la courbe est très resserrée ; sinon elle est plus ample. C'est ce qu'à la fin du XVIII^e siècle, les savants appellent la « courbe de facilité » des erreurs de l'instrument en question. Ce n'est pas vraiment une courbe de Gauss, ça peut même être à peu près n'importe quoi de mathématique.

Figure 7. Une courbe n'importe comment, ou plusieurs courbes de formes diverses, demi-cercle, triangle isocèle, rectangle, courbe en U, etc. ou alors reprendre le thème du boa avalant un chameau, un mouton, un éléphant, etc.

Évidemment les savants les plus rationnels ont proposé des courbes de facilité particulières pour des raisons tout à fait convaincantes, Mais quelle que fût cette forme, aussi simple fût-elle, une courbe plate sur un intervalle donné, par exemple, les calculs qu'il y avait lieu de faire pour déterminer, à l'intérieur de la théorie, les valeurs les plus vraies possibles et leurs degrés de précision, étaient, comme dans le cas du jeu de pile ou face, très vite inextricables, et d'une inextricabilité en quelle que sorte supérieure. De quels calculs s'agissait-il ?

On a fait, mettons, 100 mesures avec le même instrument, dont on suppose connue la courbe de facilité des erreurs, par exemple, une courbe plate entre 0 et 4 : tous les nombres de l'intervalle (0,4) ont même chance d'être obtenus.

Figure 8. La densité de la loi uniforme sur (0,4) et un histogramme approximatif de cent mesures uniformes, l'un sur l'autre.

Pour tenter de réduire les imprécisions inévitables des mesures, depuis très longtemps et en particulier au XVIII^e siècle, on en fait la moyenne arithmétique. Pourquoi ? C'est comme ça qu'on a toujours fait et visiblement ça marche assez bien. Mais de combien a-t-on amélioré la précision, si tant est qu'on l'ait améliorée ? La moyenne n'est-elle pas, dans certains cas, pire que l'une des mesures faite avec un soin particulier² ?

Peut-on préciser, comme nous l'avons fait tout à l'heure pour le jeu de pile ou face ? Oui sans doute, il suffit de calculer la loi de facilité de la moyenne arithmétique des 100 mesures et regarder si elle est vraiment plus resserrée que la loi d'une seule des mesures, et si possible de combien elle est plus resserrée. Or ce calcul, personne n'est parvenu à le faire au XVIII^e siècle, pas même Euler et Lagrange, les maîtres incontestés des mathématiques du temps. Et Laplace, né en 1749 à Baumont-en-Auge, qui les défie sur tous les sujets possibles, n'y est pas parvenu non plus, dans ses travaux des années 1770-1780, lui qui a entrepris de résoudre toutes les difficultés analytiques de la mécanique céleste et de la théorie des chances.

Que faire en attendant ? Oublier le problème provisoirement et faire autre chose, autrement. En particulier chercher à résoudre des problèmes plus simples : 100 mesures étant données de loi de facilité connue, quelle valeur, tirée de ces mesures, possède telle ou telle propriété mathématique remarquable (en espérant que celle-ci ait un rapport quelconque avec la réalité des choses) ; c'est ce qu'ont fait Euler, Daniel Bernoulli, Laplace, Legendre et même le jeune Gauss à peine âgé de vingt ans. Legendre et Gauss, pour leur part, proposent d'utiliser le critère des « moindres carrés » (on ne sait pas qui des deux en a eu le premier l'idée, parce que c'est Legendre qui a publié la méthode le premier en 1805, mais Gauss affirme qu'il la connaissait et qu'il l'utilisait longtemps auparavant). Dans le cas le plus simple que nous examinons, la méthode des moindres carrés consiste à choisir pour mesure vraie, le nombre m qui rend minimum la somme des carrés des erreurs observées. Pour simplifier, considérons seulement trois mesures d'une même constante naturelle, faite avec le même instrument. On a obtenu mettons 1, 2 et 3, la méthode des moindres carrés consiste à choisir comme valeur de la constante en question le nombre x qui rend la plus petite possible l'expression :

² Pour certaines distributions de probabilités, cela peut en effet arriver. Mais il s'agit, selon la formulation de P. Lévy, de « lois exceptionnelles », que l'on ne rencontre que très rarement dans les applications du calcul des probabilités.

$$(x-1)^2 + (x-2)^2 + (x-3)^2$$

Pour rendre minimum une telle expression toujours positive, il suffit de chercher où s'annule sa dérivée. On obtient ainsi :

$$x = \frac{1+2+3}{3} = 2$$

c'est-à-dire la moyenne arithmétique des trois mesures. Si bien que la méthode ancestrale de la moyenne arithmétique se trouve, d'une certaine façon, justifiée mathématiquement : elle se confond dans le cas le plus simple avec la méthode du prince des mathématiciens, Gauss, et de Legendre (qui n'est pas prince mais au moins baron). Inversement la méthode de ces deux illustres savants se trouve justifiée par une pratique centenaire. La difficulté évidente de ce raisonnement, c'est que rien ne prouve que ces deux méthodes qui s'accordent si bien ensemble, s'accordent avec le monde extérieur et fournissent une valeur approchée de la véritable constante. Il faudrait qu'elles nous disent, au moins, de combien elles se trompent probablement, lorsqu'elles se trompent. Mais cela suppose que l'on connaisse la loi de facilité des erreurs et qu'on sache calculer à partir d'elle la loi de facilité de la moyenne arithmétique ou bien la loi de facilité du résultat fourni par la méthode des moindres carrés. Or on ignore généralement la première et, de toute façon, on ne sait pas calculer la seconde.

En 1809, Gauss publie sa première théorie de la méthode des moindres carrés. Elle est purement gaussienne (les courbes de Gauss résolvent tous les problèmes). Gauss observe en effet que la courbe qui ne s'appelle pas encore de Gauss rend la méthode des moindres carrés entièrement cohérente et qu'elle est essentiellement la seule. Si l'on suppose que les mesures sont distribuées suivant une courbe de Gauss, la valeur la plus vraisemblable de la constante cherchée est celle que donne la méthode des moindres carrés et c'est la moyenne arithmétique, qui, elle-même, se trouve avoir également une loi de facilité du type de Gauss. De sorte qu'on peut répondre à toutes les questions évoquées ci-dessus. La difficulté de cette théorie est qu'elle est circulaire. Cette fois-ci, la méthode des moindres carrés et la courbe de Gauss se confortent l'une l'autre, mais comme rien ne conforte l'adéquation de l'une ou de l'autre à la réalité des choses, rien n'autorise à en déduire quoi que ce soit sur la validité de l'ensemble, qui n'est, tout au plus, qu'un exercice de style mathématique. C'est du moins l'objection que lui fit aussitôt Laplace, qui appréciait assez peu que le jeune directeur de l'Observatoire de Göttingen vienne piétiner ses plates-bandes. Laplace n'avait pas tort sur le fond et Gauss dut en convenir intérieurement. Toutefois, il semble bien que ce soit Gauss qui, le premier, ait proposé de choisir la courbe de Gauss comme « courbe d'erreur » de référence et qui ait compris que cette « hypothèse gaussienne » permettait de faire tous les calculs d'erreurs, dont l'astronomie de précision et la géodésie nouvelle avaient besoin. La statistique gaussienne est mathématiquement correcte, quand bien même elle n'aurait qu'un rapport indéterminé avec les réalités qu'elle est censée décrire.

Figure 9. Gauss avec son bonnet rond sur la tête, devant une courbe en cloche en forme de danseuse orientale, qui dit :
« I love the courbe de Gauss », ou la même chose en allemand,
bien que ça sonne moins bien

C'est là la seule connection véritable de Gauss avec la courbe de Gauss, et il n'en était guère satisfait puisqu'il publiera douze ans plus tard une seconde théorie de la méthode des moindres carrés, beaucoup plus belle et riche que la précédente, où l'on ne fait plus du tout l'hypothèse que les mesures sont distribuées suivant la courbe de Gauss, et qui est encore enseignée actuellement, dans toutes les universités du monde, sous le nom de théorie de Gauss-Markov, ce dernier l'ayant retrouvée indépendamment au début du XX^e siècle.

Donc Gauss, dans les années 1820, répudie la courbe de Gauss, qui, vexée sans doute, pour se venger, intrigua de façon qu'on l'associe à jamais à son ex-prétendant, qui l'avait si cruellement rejetée après l'avoir un court moment regardée.

Mais, de son côté, la courbe de Gauss avait trouvé à se consoler, et allait atteindre enfin la reconnaissance universelle à laquelle elle aspirait depuis Moivre.

Entre temps, en effet, et pour de toutes autres raisons, Laplace a repris le problème qui l'avait arrêté en 1775, puis en 1785, puis en 1795, ..., et qu'il avait gardé soigneusement en réserve. Nous avons énoncé plus haut ce problème, exactement dans les termes laplaciens, on fait 100 mesures dont les courbes de facilité sont plates entre 0 et 4. Quelle est la loi de facilité, la courbe d'erreur, de leur moyenne arithmétique, ou de leur somme ?

En 1810, Laplace, dont l'obstination et la force analytique sont légendaires, parvient enfin à une solution satisfaisante de ce problème, apparemment anodin, mais qui conditionne de fait toute la théorie des erreurs de mesures, c'est-à-dire toutes les sciences qui utilisent des nombres, c'est-à-dire, disent certains, toutes les sciences.

La courbe de facilité de la moyenne arithmétique des cent mesures est à fort peu près, elle-même, une courbe de Gauss dont l'axe de symétrie est située à l'abscisse 2 et l'écart-type, en l'occurrence 0,115, se calcule très simplement. Si, au lieu de cent mesures, nous en avons n (n assez grand, mettons supérieur à 10), la courbe de facilité de la moyenne arithmétique serait encore proche d'une courbe de Gauss centrée à l'abscisse 2 et d'écart-type $\frac{2}{\sqrt{3n}}$, et l'approximation est très bonne et d'autant meilleure que n est grand.

Figure 10. Une courbe plate entre 0 et 4, superposée à deux ou trois cloches d'écart-type $\frac{2}{\sqrt{3n}}$

Plus fort encore : si on remplace la courbe plate que l'on s'est donnée par n'importe quelle courbe avec autant de creux et de bosses qu'on voudra, il en sera de même, la somme ou la moyenne arithmétique d'un grand nombre de mesures distribuées de cette façon, aussi bizarroïde soit-elle, a une courbe de facilité du type de Gauss dont les paramètres se calculent aisément à partir des mesures déjà faites. C'est le « théorème de Laplace », dont les applications sont universelles, puisqu'elles ne supposent quasiment rien³ sur les courbes de facilité des données, et qui est, depuis, l'un

³ Mathématiquement, il faut quelques conditions précises, et très généralement satisfaites ; il s'agit-là d'une question de technique mathématique qui nous éloignerait trop de notre propos.

des théorèmes les plus fondamentaux de la théorie des probabilités, telle qu'on l'enseigne partout dans le monde. La démonstration de Laplace n'est pas simple, c'est le moins qu'on puisse dire. Elle utilise la méthode de Laplace que nous avons évoqué plus haut à propos de la formule de Stirling et la théorie, dite des intégrales de Fourier, dont Laplace est, comme on sait, le véritable créateur, avec, en compléments, quelques appels énigmatiques à la généralité de l'analyse, qui la rendront pratiquement illisible longtemps. Le théorème de Laplace, reproduit dans la *Théorie analytique des probabilités*, que Laplace a écrit en 1812, pour lui servir d'écrin, ne sera véritablement accepté par les mathématiciens, et utilisé à bon escient par les statisticiens, qu'un siècle plus tard et non sans mal, mais ceci est une autre histoire. Ceux qui tenteront de le lire dans la première moitié du XIX^e siècle (Gauss est l'un d'eux, mais aussi Fourier, Poisson, Cauchy, Dirichlet et quelques autres) reconnaîtront avec Augustus de Morgan qu'il s'agit là du « Mont-Blanc de l'analyse mathématique ». La courbe de Gauss domine enfin l'Europe mathématique, même si personne ou presque ne s'en est aperçu.

Figure 11. Une courbe en cloche en forme de Mont-Blanc, et un petit Laplace en costume de chancelier du Sénat, avec un piolet et un sac à dos, qui dit : « j'ai vaincu la courbe de Gauss », ou quelque chose comme ça

Quoi qu'il en soit, la courbe de Gauss prend sa véritable dimension mathématique en 1810. Dès qu'on additionne des observations numériques soumises à des erreurs accidentelles de même genre, la courbe de facilité résultante est mathématiquement du type de Gauss. On peut alors calculer des fourchettes où se trouvent situées les véritables constantes de la nature avec une très grande probabilité, comme on calculait avec Moivre les nombres de piles les plus vraisemblables, en un nombre conséquent de jets d'une pièce. Dans les deux cas, les intervalles et les probabilités sont donnés par la même table, la table de la courbe de Gauss.

Seulement la table (établie par Kramp à la demande de Laplace, dès la fin du XVIII^e siècle), parce que la courbe de Gauss n'existe pas encore en tant que courbe. Loin de s'exhiber dans les magazines, comme elle le fait actuellement, elle se cache alors pudiquement sous la table. Nulle part on ne la voit dessinée, ni dans les œuvres de Gauss, ni dans celles de Laplace, ni dans celles de leurs contemporains. D'ailleurs les courbes de facilité ne se montrent pas, ça ne se fait pas. On imagine bien qu'elles doivent se trouver dans les marges des brouillons des grands hommes qui les manipulent, mais ce ne sont que des graffitis furtifs destinés à la rêverie puis à la poubelle.

INTERMÈDE

Finalement, si on a bien suivi, pour le mathématicien, la courbe de Gauss intervient dès qu'il fait un calcul de chances un tant soit peu composé. C'est très bien. Bravo le mathématicien ! Mais cela n'explique pas pourquoi on retrouve la courbe de Gauss dans les endroits les plus inattendus, les plus prosaïques, où l'on ne se préoccupe aucunement de calcul des chances.

Figure 12. Il faudrait un graphisme qui illustre la différence des points de vue d'un mathématicien et d'un non-mathématicien. Quelque chose comme le savant Cosinus debout devant son tableau noir où serait tracée une courbe de Gauss, et Bécassine avec son panier rempli de courbes de Gauss en bottes, comme des poireaux. Mais bien sûr en moins troisième âge, avec des personnages de BD ou de films actuels, que je ne connais pas, et si possible en inversant les sexes (pour contourner la censure)

Il faut donc regarder la courbe de Gauss d'une autre façon que le mathématicien. Voyons cela.

D'abord, les astronomes et les géodésiens de terrain, pour qui avaient travaillé les grands mathématiciens que nous venons de quitter, ont fait, on s'en doute, ce qu'il était naturel de faire à ce point de l'histoire. Dans les années 1830, ils ont commencé à dresser des tableaux statistiques de mesures. Le premier à l'avoir fait vraiment, du moins le premier qui nous l'ai dit, c'est Bessel, un disciple et un correspondant de Gauss. Bessel travaillait à Königsberg et Gauss à Göttingen. Et en effet, Bessel observe que les mesures d'on ne sait quelles ascensions droites, répétées suffisamment, ont un tableau de fréquences qui ressemble beaucoup aux tables de la courbe de Gauss :

Figure 13. Les données de Bessel, on les trouve dans le livre de Stephen Stigler, « The History of Statistics », 1986, p. 204

L'hypothèse gaussienne n'est donc pas si absurde. Laplace est mort en 1827, il ne peut en convenir, mais, dans ses derniers travaux géodésiques des années 1820 (il a alors 75 ans), guidé par son théorème, il admet déjà dans ses calculs l'hypothèse gaussienne, sans le crier sur les toits, pour décrire par exemple les variations du coefficient de réfraction et d'autres quantités anarchiques qui résultent de l'accumulation des contributions des multiples facteurs en jeu. Le théorème de Laplace suggère d'ailleurs que tous les phénomènes résultant de l'addition de petites causes fortuites et peu dépendantes les unes des autres seront du type gaussien. Bessel s'est donné la peine de le vérifier. Dorénavant on apprendra, aux astronomes et aux géodésiens débutants, le calcul d'erreurs gaussien.

On nous objectera avec justesse qu'on reste encore dans la sphère savante, qu'il s'agit de mesures de précisions qui n'affectent que marginalement la vie quotidienne des populations. Certes, mais on doit convenir cependant que, timidement, à petits pas, la théorie de Gauss-Laplace, jusque-là dissimulée dans des traités hautement impénétrables, se vulgarise quelque peu, assez, en tout cas, pour attirer l'attention de quelques-uns de ceux que l'on commence à appeler les « statisticiens », un néologisme du début du XIX^e siècle, des gens qui, par passion ou par nécessité professionnelle, accumulent des chiffres sur tous les sujets possibles, populations, productions agricoles ou industrielles, activités de toutes sortes, enregistrements météorologiques, etc. Ces statisticiens ne sont pas des mathématiciens professionnels, et les travaux de Gauss ou ceux de Laplace, leur sont inaccessibles. On en trouve même, et de très influents, qui sont tout à fait allergiques aux mathématiques de leur temps.

Certains, pas tous, commencent à publier des « tableaux » statistiques, des listes de chiffres qui disent combien de fois telle éventualité s'est produite dans l'ensemble de leurs observations. Ils ne tracent généralement pas de « courbes représentatives » de

leurs tableaux. Il faudrait pour cela qu'ils reconnaissent ou qu'ils attribuent à leurs tableaux, à leurs chiffres fluctuants et erratiques, une sorte d'existence autonome, une personnalité propre, une autonomie, une indépendance, une constance particulière, abstraction d'un degré supérieur qu'ils n'envisagent pas en général. Ce qui n'empêche pas certains d'adorer les représentations graphiques, les cartes colorées, les courbes sinueuses décrivant les variations d'une quantité particulière au cours du temps, productions de blé ou températures par exemple, les courbes de mortalité qui donnent le nombre de survivants d'une population donnée au bout de tant d'années, etc., toutes espèces de courbes qui peuvent être vues comme des abstractions mathématiques, dès lors que nos statisticiens lissent un peu leurs données ou les corrigent pour que leurs dessins soient plus lisibles. Pour se justifier, les plus avisés d'entre eux se réclament d'un principe général, issu de la sagesse populaire ou du théorème de Bernoulli : prises sur le moyen ou le long terme, les choses restent à peu près les mêmes. Par exemple, la quantité d'eau de pluie qui tombe en un lieu donné est relativement constante d'une année sur l'autre, il y a bien sûr des années sèches et des années pluvieuses, mais les variations fortes sont rares. Et ce principe est encore plus vrai, si, au lieu des nombres absolus, on considère des moyennes et des proportions, pour lesquelles on observe quasiment plus de variations du tout d'un tableau à l'autre. Par exemple, la proportion des naissances de garçons aux naissances annuelles est toujours légèrement supérieure à $1/2$. (Ce fait, connu de longue date, est d'ailleurs le premier qui ait été précisé à l'aide du théorème de Bernoulli, comme l'a indiqué, dès 1713, Nicolas Bernoulli, un neveu de Jacques). La proportion des naissances à la population totale est-elle aussi constante (et égale à la durée moyenne de vie, et cette fois-ci c'est Laplace qui a tenté d'appliquer sa théorie à un tel phénomène), etc. Plus curieux encore, on observe que ce même principe s'applique à des faits de l'ordre moral. La proportion des crimes rapportés à la population totale est constante d'une année sur l'autre, celle des suicides également, la proportion des lettres mises au rebut pour défaut d'adresse par rapport au nombre total de lettres, également, et ainsi de suite. Il y a, un peu partout, des rapports constants, et au sein même des phénomènes qui semblent ne dépendre que du libre-arbitre de chacun, de ses choix personnels. Voilà qui est intéressant : à défaut de courbes constantes qu'ils ne voient pas encore, les statisticiens se plaisent à mettre en évidence des constantes remarquables au milieu des variations désordonnées de leurs chiffres et de leurs tableaux. Dès lors, c'est une course aux moyennes et aux proportions fixes (ou supposées telles), avec une préférence marquée pour les plus inattendues, les plus surprenantes.

Mais pas la moindre courbe de fréquences en vue, encore moins de courbe de Gauss qu'au reste, personne n'a encore nommée. C'est dire si elle n'existe pas ! Patience.

Le statisticien le plus connu du XIX^e siècle, le seul dont nous allons parler, s'appelle Adolphe Quetelet. Il est né en 1796 à Gand et a reçu une bonne éducation scientifique. Très jeune, il a fait d'intéressants travaux de géométrie, d'astronomie, de météorologie, etc. avant d'aborder les questions de populations et de statistiques qui nous intéressent ici. C'est donc un savant intermédiaire, une espèce rare et généralement mal vue, mais qui a eu la chance de naître belge et d'être en position de responsabilité au moment où la Belgique aspire à l'indépendance et la conquiert en 1830. Il peut ainsi faire œuvre originale de statisticien au sein de l'Académie de Bruxelles et des principaux organismes universitaires belges, avec les moyens de diffusion et la reconnaissance qu'il n'aurait assurément pas eus en d'autres temps ou ailleurs. Quetelet astronome connaît les courbes de facilité des erreurs. Il a certainement entendu parler du théorème de Laplace et a lu avec enthousiasme l'*Essai philosophique sur les*

probabilités, où Laplace expose ses résultats pour un public plus large, et donne au principe général des statisticiens une sorte de consistance analytique, en étalant de part et d'autre de toute moyenne constante des dégradés symétriques qui suivent les tables de la courbe de Gauss (toujours anonyme).

C'est la chance historique de la courbe de Gauss. Elle va s'en saisir aussitôt et jeter son dévolu sur Quetelet, qui succombera vite, on s'en doute bien. Et avec lui, inéluctablement, tous les statisticiens, c'est-à-dire tout le monde.

Figure 14. Une courbe en femme fatale qui emporte sous son bras le bel Adolphe en bas de soie et culotte à l'ancienne, devant le Manneken Piss. On peut évidemment faire bien pire, par exemple, une courbe de Gauss soulevant un pan de sa jupe, qui hèle le pauvre Adolphe, dans le genre : « tu viens chéri, y-a des frites », etc.

En 1846, Quetelet, après vingt ans de travaux statistiques divers, publie, à son tour, une version grand public de la philosophie laplacienne, ou bernoullienne, enrichie d'exemples très remarquables empruntés aux statistiques du temps, qu'il connaît mieux que personne. Ce sont les *Lettres à S. A. R. le grand-duc régnant de Saxe-Cobourg et Gotha sur la théorie des probabilités appliquée aux sciences morales et politiques*.

Les *Lettres* commencent par exposer sans aucun formalisme la théorie du jeu de pile ou face, dont nous avons parlé. Pour l'une des premières fois, on voit apparaître dans un ouvrage imprimé, une véritable courbe en cloche. Quetelet résume, en effet, son propos par la courbe suivante qui a de toutes petites marches d'escalier, mais qui possède indéniablement la bonne forme, et qui se passe de commentaires :

Figure 15. La première courbe de Quetelet (p. 396) est reproduite dans le livre de Stephen Stigler cité Figure 13, p. 209

Quetelet introduit ensuite la fonction de Gauss et la trace en plusieurs endroits. Le grand-duc de Saxe-Cobourg et Gotha peut admirer enfin la courbe de Gauss.

Quetelet aborde ensuite les applications. L'une des plus curieuses consiste à étudier un tableau, publié par l'armée britannique vers 1840, donnant les tours de poitrine d'environ 4000 soldats écossais. Sans doute une commande de l'intendance militaire en vue de la confection de tenues aux bonnes tailles (Freud aurait vu là matière à développements, mais nous n'insisterons pas sur ce point)

Plutôt que se lancer dans de longs commentaires, Quetelet a l'idée de comparer le tableau des tours de poitrine à celui de la loi de Gauss correspondante. Voilà le résultat :

Figure 16. Le tableau des tours de poitrine, avec en surimpression un Écossais moyen en kilt et dans un coin, Nessie, le petit monstre du Loch Ness qui dit : « I love the courbe de Gauss », etc.

Difficile de faire mieux.

Quetelet ne superpose pas la courbe de fréquence des tours de poitrine à la courbe de Gauss ajustée sur le tableau. Cela ne se fait pas encore. Les courbes de fréquences, les « histogrammes », n'apparaîtront systématiquement qu'à la fin du siècle, avec la statistique britannique. Le mot anglais « histogram » date de 1895, sa traduction française normalisée de 1948.

La loi de Gauss régit les tours de poitrine écossais ! On savait déjà que, dans des données de ce genre, se cache toujours une moyenne, une proportion constante, reproductible par d'autres tableaux de données semblables. Mais là, c'est encore plus beau, une cloche, une vraie cloche, dont on imagine bien qu'elle va se retrouver à l'identique, si l'on prend un second groupe de 4000 écossais en âge de porter les armes. On imagine aussi, et l'on imaginera effectivement très vite, que les tours de poitrine n'ont rien de particulier à cet égard et qu'il en sera de même pour les tours de mollets, les tours de taille, les périmètres crâniens et tout ce qu'on voudra se donner la peine de mesurer sur des individus de même type, écossais ou non, militaires ou non. Il ne faudrait en conclure que tout est dit. Peu à peu, on mettra en avant d'autres « courbes » qui copient fort bien d'autres types de données statistiques : la loi de Poisson dans le cas des événements rares, celle de Pareto pour les distributions de revenus, la loi exponentielle et ses dérivés pour les durées de vie, etc. Mais cela ne se fera que progressivement à la fin du XIX^e siècle et nous nous éloignons de notre sujet.

Comment Quetelet explique-t-il l'identité de la courbe du jeu de pile ou face répété longtemps et celui des tours de poitrine écossais ?

Fort simplement, à la manière de Laplace. Jouer cent fois à pile ou face par exemple, c'est accumuler cent petits hasards indépendants. On doit donc obtenir, à fort peu près, une courbe de Gauss : c'est le théorème de Moivre (Laplace). Le tour de poitrine d'un soldat écossais, c'est le résultat de l'accumulation de cent ou de mille petites circonstances particulières, le lait de sa nourrice, son coup de fourchette, la qualité de son whisky, ses activités physiques, et pourquoi pas les caractéristiques de tous ses ascendants maternels et paternels, etc. Il n'est donc pas très étonnant que nous retrouvions une courbe en cloche distribuée autour du tour de poitrine moyen, qui est une sorte de caractéristique raciale des jeunes Écossais : c'est encore le théorème de Laplace.

Bien sûr, et certains de ses collègues ne manqueront pas de le faire observer, on peut ajuster à la cloche des soldats écossais bien d'autres fonctions mathématiques plus simples que la fonction exponentielle de Gauss. Mais on y perdrait beaucoup au point de vue théorique. Comment ignorer le Mont-Blanc lorsqu'il est là devant nous, et aussi au point de vue pratique, la courbe de Gauss, avec ses deux paramètres caractéristiques et sa table toute faite, se plie le plus simplement du monde à nombre de données possibles.

L'explication de Quetelet par les causes multiples accumulées, convenablement extrapolée, amène tout naturellement à la « théorie de l'homme moyen ». Selon Quetelet, dans une population donnée, il y a toujours un type générique, soumis à des variations accidentelles ou circonstanciées, les hasards multiples de la vie, et ces variations ne sont pas quelconques, elles se distribuent suivant une courbe, la courbe de Gauss, que Quetelet appelle la « loi des causes accidentelles ».

La loi de Gauss est enfin nommée. Pas du nom de Gauss (qui était encore vivant en 1846 et qui aurait certainement protesté), mais c'est égal, elle existe et elle n'en finira plus d'exister (la preuve !).

Dans la seconde moitié du XIX^e siècle, la loi des causes accidentelles va connaître un engouement certain. Les statisticiens qui se livraient jusqu'alors au recensement méticuleux des rapports constants, vont maintenant partir à la recherche des cloches constantes. Ils en découvriront partout, et même un peu trop, mais ce sont des erreurs de jeunesse et seul l'enthousiasme compte !

Pour ne donner qu'un seul exemple. Racontons l'histoire du petit Bartel.

Peut-être avez-vous entendu parler de Bartel Leendert van der Waerden. C'est un mathématicien important du XX^e siècle, né à Amsterdam en 1903, l'un des plus abstraits qui soit. Ses travaux en algèbre générale, en géométrie algébrique, etc. sont très connus, et si Gauss y intervient ce n'est assurément pas par l'intermédiaire de « sa » courbe. Pourtant dans sa liste impressionnante de publications, on trouve un court traité de statistique. Ce n'est visiblement pas une œuvre de commande mais l'occasion d'un témoignage de piété filiale. Lorsqu'il était enfant, raconte-t-il dans ses souvenirs, son père l'avait accompagné en forêt et lui avait demandé de ramasser les feuilles d'un bel arbre qui se trouvait là. Ce devait être l'automne. Rentrés à la maison, son père lui avait dit de ranger soigneusement les feuilles les unes à côté des autres par ordre de taille, de la plus petite à la plus grande, de sorte que les bas des feuilles soient alignés. Il lui avait ensuite fait observer la forme en s allongé qu'adoptaient les pointes de feuilles.

Voici, lui avait-il dit, la courbe de Quetelet qui permet de voir que les individus moyens sont les plus nombreux

(Cf. B.L. van der Waerden, *Statistique mathématique*, éd. française, Paris, Dunot, 1967, p. 66)

C'était là, on s'en doute, une autre façon de représenter la courbe de Gauss, celle qui consiste à porter en abscisse le nombre cumulé des feuilles de taille inférieure à une taille donnée en ordonnée. Mais il n'est pas nécessaire d'en dire autant. Il suffit de regarder :

Figure 17. La courbe en s des longueurs des feuilles d'un châtaignier, par exemple. Avec les feuilles bien rangées les unes à côté des autres

ÉPILOGUE ÉPONYMIQUE

Une question angoissante court tout au long de ce texte minuscule : par quels hasards multipliés le nom de Gauss s'est-il trouvé définitivement associé dans la culture mondiale à la courbe de Gauss ? Nous avons laissé cette dernière vers le milieu du XIX^e siècle sans nom ou plutôt avec une abondance de noms divers, loi des erreurs accidentelles, courbe des erreurs, courbe de Quetelet, ..., que la postérité n'a pas retenus, pour adopter définitivement le nom de Gauss. Que s'est-il donc passé ? Il s'agit là en fait d'un problème général, le problème des éponymes, qui a été fort bien analysé au chapitre 14 d'un livre important de S. M. Stigler, *Statistics on the Table*, Harvard University Press, 1999. Stephen Stigler traite particulièrement l'exemple de la « loi de Gauss », et dans tous les détails souhaitables, nous y renverrons donc le lecteur.

Rappelons-en les termes rapidement. Le savant illustre qui donne son nom à quelque chose n'est jamais l'inventeur de la chose en question, ou si peu que ce n'est pas la peine d'en parler. Lorsqu'on affecte un nom à un énoncé scientifique dont l'importance est universellement reconnue et dont l'histoire est mal connue, comme le

sont toutes les histoires, tout se passe comme s'il s'agissait de choisir une sorte de héros tutélaire qui peut-être eut à connaître l'énoncé dont il s'agit, mais que l'on entend d'abord honorer et qui honore, par contre coup, et magnifie la chose nommée, ici, en l'espèce, la courbe en cloche. D'ailleurs les Grecs ne l'entendaient pas autrement qui ont nommé Athènes pour honorer Athéna mais aussi pour qu'elle protège leur cité. La recherche des origines éponymiques se trouve ainsi compliquée considérablement, puisqu'elle cumule les difficultés et les aléas de toute recherche d'origine et de paternité intellectuelle (recherche d'ailleurs mal vue des véritables historiens mais très prisée des amateurs, comme vous ou moi), avec les complexités sans nombre de l'histoire des mentalités et des passions scientifiques. Ceci étant dit en préambule, la question se pose et demeure : pourquoi diable est-ce Gauss qui a été retenu par l'usage, à partir de la seconde moitié du XX^e siècle, plutôt que Moivre ou Laplace, ou Quetelet ou mille autres qui avaient plus de titres que lui à ce rôle ? Stigler donne plusieurs pistes. D'abord les géodésiens de langue allemande, les premiers à avoir fait un usage intensif de la méthode de Gauss pour résoudre les grandes triangulations géodésiques, méthode dans laquelle la courbe des erreurs intervient en effet. D'où un glissement naturel, à partir des années 1870-1880, de la méthode de Gauss, qui est véritablement l'œuvre de Gauss, à la courbe de Gauss qui ne lui doit guère, et une première éponymie géodésique « allemande » pour l'essentiel. Mais, selon Stigler, ce n'est assurément pas une raison suffisante. Elle pêche par excès de spécialisation, de marginalité et de nationalisme. Il faut chercher plus loin. Le premier auteur « généraliste », qui paraît avoir associé la courbe de Gauss à Gauss, est un savant français généralement mal considéré des historiens des mathématiques mais qui a joué un rôle considérable au sein de la communauté scientifique française dans la seconde moitié du XIX^e siècle, Joseph Bertrand, auteur d'un célèbre traité de calcul des probabilités (Paris, Gauthier-Villars, 1888 et 1889). Publié peu de temps avant que la théorie des probabilités ne fasse un retour en force dans la science mondiale, le livre de Bertrand, le seul disponible alors à bien peu près, et, qui plus est, d'une grande clarté pédagogique, a été beaucoup lu et copié en Europe dans les années qui ont suivi. De sorte que la « loi de Gauss », qui y est étudiée longuement, se retrouve dans nombre de travaux probabilistes du début du XX^e siècle, ceux de Bortkiewicz, Poincaré ou Lévy par exemple. On comprend que la candidature gaussienne ait été favorisée par de tels patronages, d'autant qu'il y avait urgence. La statistique mathématique va se répandre en effet assez rapidement dans l'entre-deux-guerres et plus rapidement encore après la seconde guerre mondiale. Il faut par conséquent des enseignements et des livres de calcul des probabilités dans lesquels évidemment la courbe de Gauss apparaît en bonne place et qu'il faut bien nommer pour qu'on l'identifie et qu'on apprenne à s'en servir.

Ainsi ce serait un savant français qui aurait été à l'origine d'une des principales dénomination qui nous intéresse ici, laquelle choqua beaucoup certains autres savants, français ou non, plus au fait de l'histoire de leur science, particulièrement Maurice Fréchet, qui lutta pied à pied pour qu'on nommât loi de Laplace, ce qui déjà, le plus souvent, s'appelait loi de Gauss. Il fit même « normaliser » la locution « loi de Laplace » par l'AFNOR en 1948, sans grand succès, même en France. C'est d'ailleurs là une règle générale de la théorie éponymique : un éponyme ne se décrète pas, il est consacré par l'usage. De plus, pour tenter d'influencer l'usage et d'imposer un nom, il vaut mieux jouer contre son camp. En l'occurrence, pourtant, Bertrand ne jouait pas contre son camp en préférant Gauss à Laplace, mais pour le sien. Ne se proposait-il pas ouvertement dans son traité de présenter la théorie des probabilités sans jamais recourir à la « théorie analytique » de Laplace, péremptoirement jugée archaïque et brumeuse, qu'il convenait de remplacer par le recours systématique à la combinatoire et à

l'algèbre, notamment l'algèbre de Gauss, qu'il portait au plus haut et dont il avait traduit en français, trente ans auparavant, les principaux textes probabilistes ? Et après tout, Gauss, le prince des mathématiciens, n'était-il pas une meilleure divinité tutélaire que Laplace ? On expliquerait ainsi que son nom l'ait emporté sur le long terme. Bertrand, en tout cas, était certain que Gauss était un bien meilleur parti que Laplace. Mais en lui donnant le nom de Gauss, il jouait un très mauvais tour à la courbe qu'il voulait honorer. Certes il lui faisait épouser un prince, mais il la privait du royaume que lui avait laissé Laplace en héritage, le théorème général de Laplace, qui ne figure nulle part dans le traité de 1888, non plus que sa méthode de démonstration, anticipation audacieuse (quoique brumeuse) de l'analyse de Fourier moderne. Heureusement, plus à l'Est, on n'avait pas oublié la théorie laplacienne de la courbe de Gauss, qui serait bientôt, entre les mains de Liapunov, de Markov et de beaucoup d'autres, remise à sa vraie place. Laplace, cependant, ne parviendrait pas davantage à imposer son nom à son résultat le plus remarquable. Le théorème de Laplace serait en effet baptisé « théorème central limite » (ou « limite central » au choix) par Polyà, vers 1920, et cette éponymie anonyme s'imposerait en dépit de son esthétique un peu trop germanique, Laplace devant se contenter, ici ou là, du cas particulier du jeu de pile ou face, entièrement dû à Bernoulli et à de Moivre, comme on le sait. Mais Laplace avait l'habitude, on l'avait privé déjà de l'éponymie bayésienne, lui qui plus que Bayes, l'obscur, le sans grade, la méritait tant (et cette fois-ci hélas ce fut Cournot et son Exposition de la théorie des chances de 1843 qui, à son corps défendant sans doute, joua la partie de Bertrand et son traité de 1888. On n'est trahi que par les siens !).

Il y a toujours une morale aux histoires immorales, Devant Bertrand, qui avait tant de talent et si peu de génie, le Prince n'épousait qu'une bergère aux charmes un peu frustes et, sans y gagner grand chose, se trouvait pris au piège des éponymes. On sait que Pythagore n'existe véritablement que par « son » théorème (connu dans le concret des maçons mésopotamiens et dans l'abstrait d'Euclide), Gauss risquait fort de n'exister plus que par une courbe dont il n'avait que faire. Aussi, Bernoulli, Moivre ou Laplace, qui sont à l'origine de toute cette histoire (au moins une origine possible, identifiée et authentique), vierges de toute éponymie trop visible, sortent grandis de cette ténébreuse affaire : ils gardent pour l'éternité leur intégrité, leur génie, leur mystère et leur anonymat !