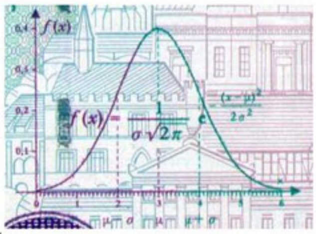
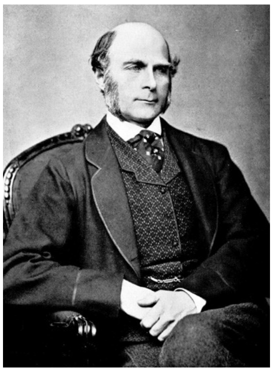
LOIS À DENSITÉ : LES LOIS NORMALES
COMPLÉMENTS



Un billet de dix Deutsche Mark, avec Carl Friedrich Gauss (1777-1855).



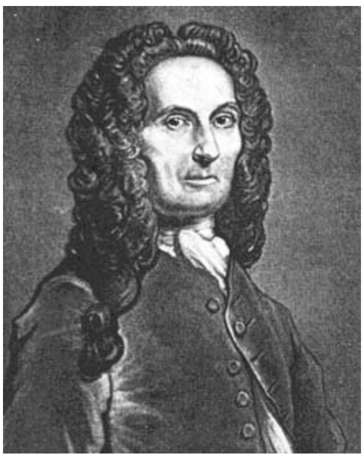
← Laplace (1749 – 1827) →



Galton (1822 – 1911)


← Quételet (1796 – 1874) →

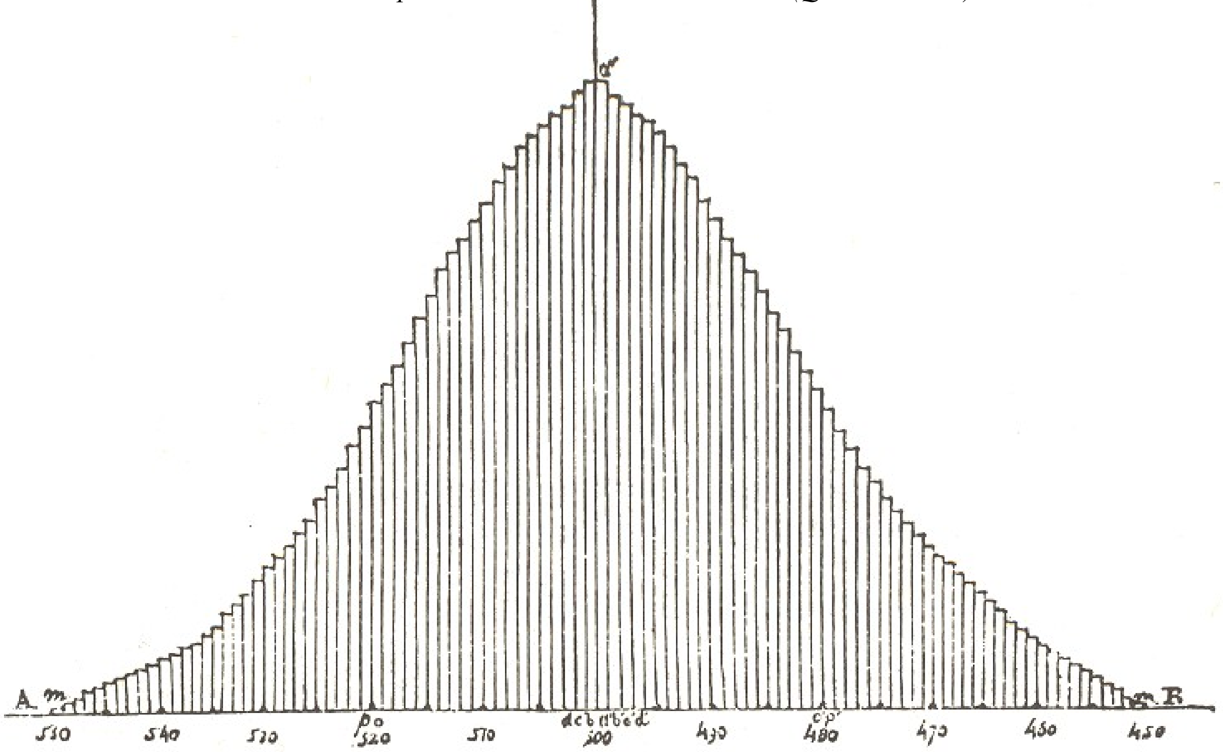
Pearson (1857 – 1936)

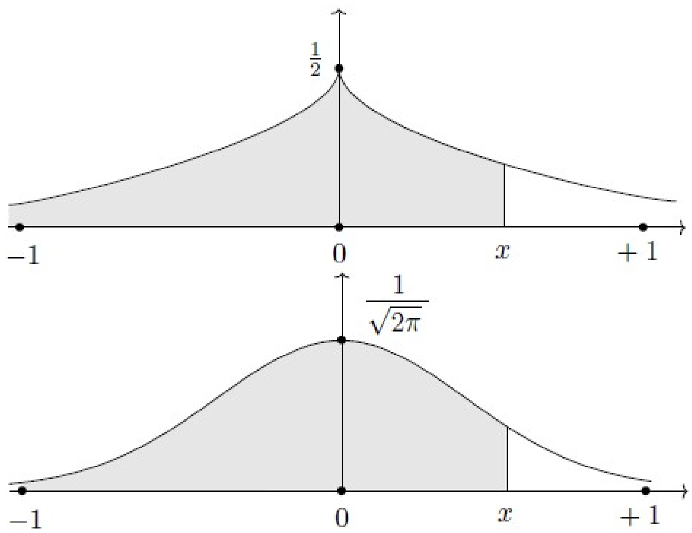


De Moivre (1667 – 1754)

Gauss (1777 – 1855)

le tour de poitrine de 5738 soldats écossais (Quételet, 1846)





Première et deuxième lois de Laplace

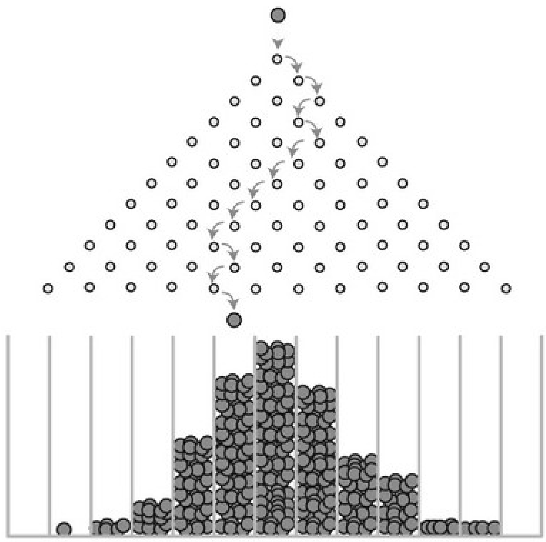


Planche de Galton

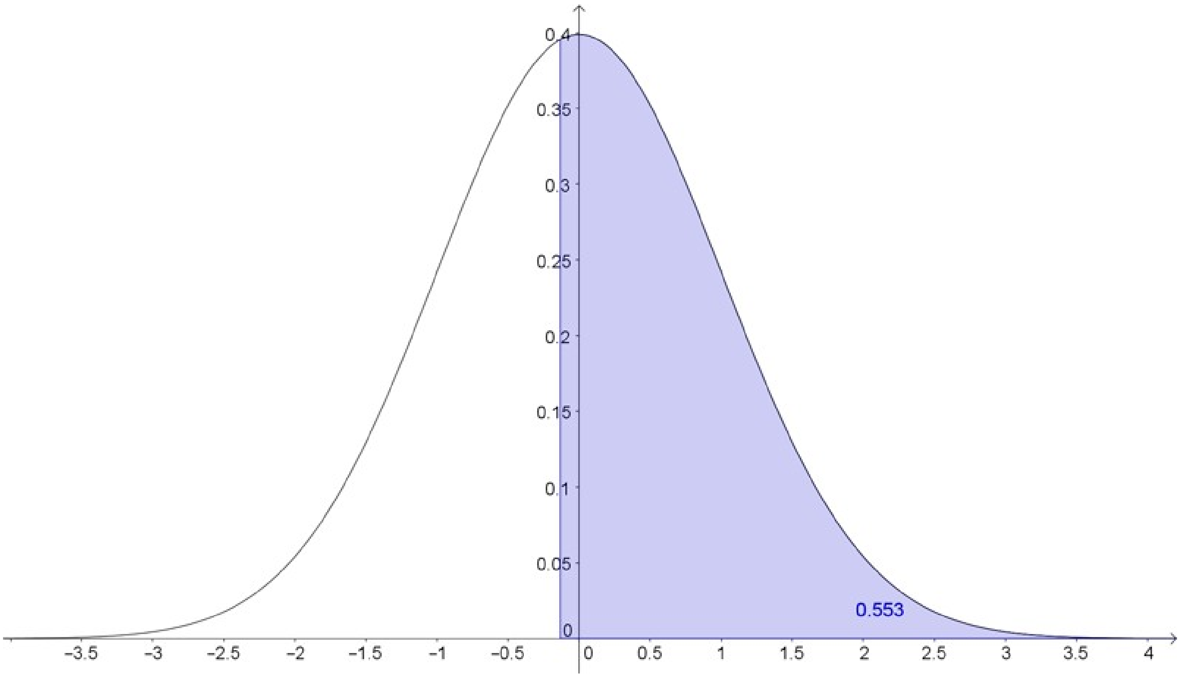
I. Loi normale centrée réduite

Propriété : si X suit une loi normale $\mathcal{N}(\mu ; \sigma)$ alors $\dfrac{X-\mu}{\sigma}$ suit une loi normale $\mathcal{N}(0 ; 1)$.

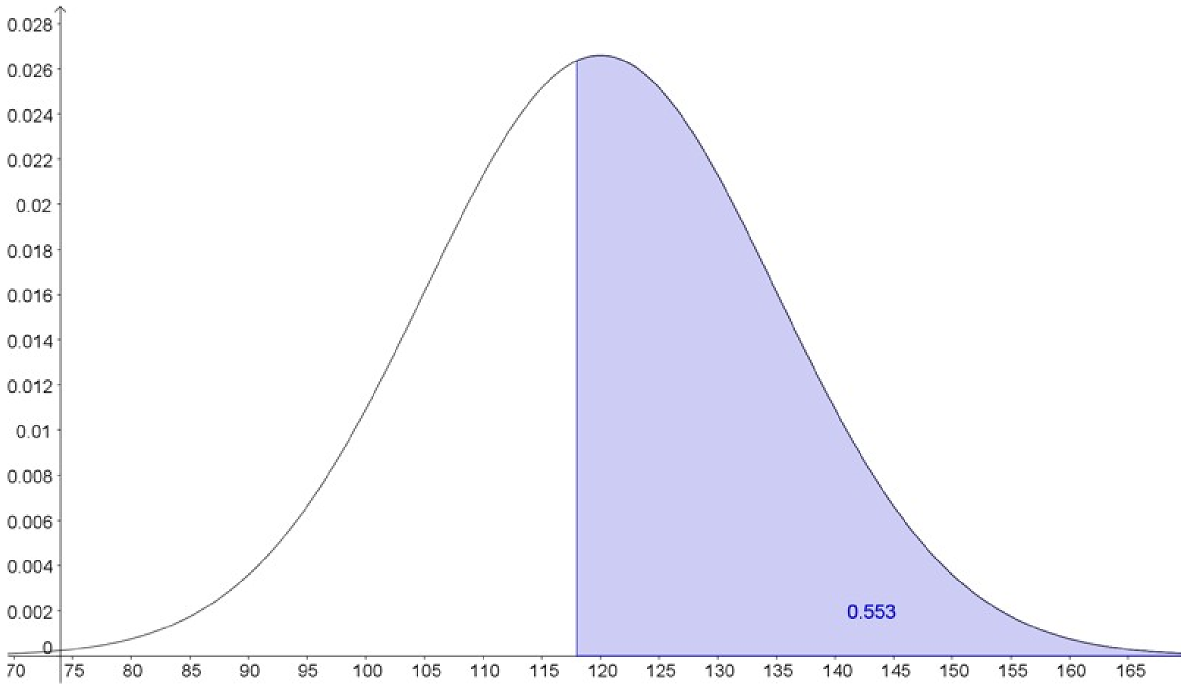
Ainsi, avant le développement des outils informatiques, on utilisait beaucoup la courbe en cloche liée à la loi normale $\mathcal{N}(0 ; 1)$, qu'on appelle ***loi normale centrée réduite***.

Par exemple, si X suit la loi $\mathcal{N}(120 ; 15)$ alors $Y=\dfrac{X-120}{15}$ suit la loi $\mathcal{N}(0 ; 1)$.

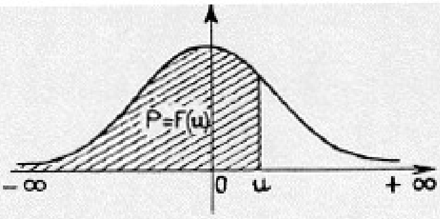
Et alors, si on devait calculer $p(X\geqslant 118)$ on écrivait : $p(X\geqslant 118)=p\left(\dfrac{X-120}{15}\geqslant\dfrac{118-120}{15}\right)\approx p(Y\geqslant -0{,}133)$ et on utilisait la courbe liée à Y (ci-dessous, à gauche) donc à $\mathcal{N}(0 ; 1)$:



ce qui revient au même que de lire graphiquement sur le dessin ci-dessus (à droite) de la $\mathcal{N}(120 ; 15)$:



Sauf qu'à l'époque la calculatrice n'existait, on utilisait des tables de la loi normale centrée réduite :



u	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7290	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9779	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

il suffisait d'utiliser la symétrie de la courbe pour écrire : $\mathrm{p}(X \geqslant -0{,}133) \approx \mathrm{p}(X \geqslant -0{,}13)$.
Or $\mathrm{p}(X \geqslant -0{,}13) = \mathrm{p}(X \leqslant 0{,}13)$ donc avec la table :

$$\mathrm{p}(X \geqslant -0{,}133) \approx \mathbf{0{,}5517} .$$

On trouvait une bonne valeur approchée du résultat que l'on obtient aujourd'hui avec la calculatrice : **0,5529**.

II. Pourquoi la loi normale est très utilisée ? Et pourquoi loi « normale » ?

Un théorème mathématique, appelé « théorème central limite » (démontré en 1809 par Pierre-Simon de Laplace) affirme que toute somme de variables aléatoires indépendantes et identiquement distribuées tend vers une variable aléatoire qui suit une loi normale. Autrement dit, ***peu importe l'expérience, si on la répète un grand nombre de fois de manière indépendante et identique, on verra apparaître une loi normale !***

Ce théorème et ses généralisations[1] offrent une explication à l'***omniprésence de la loi normale dans la nature*** : de nombreux phénomènes sont dus à l'addition d'un grand nombre de petites perturbations aléatoires, à la superposition de causes nombreuses, plus ou moins indépendantes, il en résulte que la loi normale les représente de manière raisonnablement efficace.

On dit qu'une loi normale intervient dans « la modélisation de nombreux phénomènes aléatoires possédant de nombreuses causes indépendantes dont les effets s'ajoutent, sans que l'une d'elles soit dominante ».

Par exemple, les chocs de molécules d'eau sur une molécule de pollen ou les effets des conditions atmosphériques sur le plan de vol d'un avion peuvent être modélisés par une loi normale. C'est elle aussi qui modélise les variations observées entre mesures successives d'une quantité, suite à l'erreur de mesure.

Elle permet aussi de représenter la variabilité de la glycémie à jeun, du taux de division bactérienne, etc.
Par exemple, si l'on nous dit que, dans une certaine population, chez le sujet adulte « normal » (non diabétique) la glycémie est distribuée selon une loi normale de moyenne 4,8 mmol/L et d'écart type 0,4 mmol/L, on déduit immédiatement que 95 % des sujets « normaux » de cette population ont une glycémie comprise entre 4,0 et 5,6 mmol/L…

En biométrie, un certain nombre de variables quantitatives se distribuent selon la loi normale : taille, poids, rythme cardiaque, périmètre crânien, diamètre des hématies… Les données du domaine biologique ou médical suivent souvent le modèle normal : la taille, par exemple, dépend de très nombreux facteurs : les uns héréditaires, les autres dus au milieu (alimentation, mode de vie, conditions de chauffage des habitations, etc.). Chacun de ces facteurs agissant indépendamment et pour une petite part soit vers un accroissement soit vers une diminution de la variable taille.

---

1 Ces généralisations ne nécessitent pas des lois identiques (autrement dit pas besoin que chaque variable aléatoire que l'on additionne suive la même loi) mais font appel à des conditions qui assurent qu'aucune des variables n'exerce une influence significativement plus importante que les autres.

III. Historique de la loi normale

La loi normale est associée aux noms prestigieux de Gauss et de Laplace ; dans la littérature allemande et anglo-saxonne elle est connue sous le nom de Loi de Gauss ; certains mathématiciens français y ont fait référence sous le nom de Loi de Laplace ; pour mettre tout le monde d'accord, Maurice Fréchet a proposé de l'appeler Loi de Laplace-Gauss. C'est cet usage qui a prévalu en France.

Ces mathématiciens ont tout deux introduit cette loi, mais leurs approches étaient différentes. Celle de Gauss, relative à la théorie des erreurs, aboutira à la *méthode des moindres carrés* et sera à l'origine de la théorie de l'estimation par *le maximum de vraisemblance*[2]. Celle de Laplace, d'inspiration probabiliste, prolonge les idées de de Moivre et a donné naissance à la première version du *théorème central limite*.

L'approche de Laplace

Pierre-Simon de Laplace a prolongé les recherches du mathématicien anglais de Moivre sur l'approximation de la loi binomiale par la loi normale : de Moivre avait étudié (en 1733) le comportement du nombre de pile après un nombre important de lancers de pièces <u>équilibrées</u>. En traçant les diagrammes en bâtons correspondant, il observera une « courbe en cloche », et approchera la loi binomiale par la loi normale.
En 1772, Laplace reprend donc des travaux de Bernoulli et de de Moivre et généralise leur théorème à une pièce « non parfaite » : il obtient la loi normale en tant qu'approximation de la loi binomiale en étudiant le comportement du nombre de pile après un nombre important de lancers de pièces <u>non équilibrées</u>. Là aussi apparaîtra une loi normale !

Pour établir ce théorème, Laplace inventera une méthode nouvelle qui servira beaucoup par la suite, permettant de réduire les difficultés.
L'étude de l'approximation de la loi binomiale par la loi normale n'a été pour Laplace qu'un prélude à des développements plus profonds. Il a en effet mis la loi normale en rapport avec la théorie des erreurs d'observation et c'est finalement cette approche qui s'est par la suite avérée la plus féconde pour les mathématiques, puisqu'elle aboutira au *théorème central limite*.

L'approche de Gauss

Au cours des années 1770, il ne s'agit 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 18^{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 10 fois, 100 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.

Dans ses travaux, en 1809 et 1816, Gauss s'intéresse à la loi normale pour des calculs en astronomie, notamment le mouvement des corps célestes. Pour minimiser les erreurs obtenues à partir de différentes

2 Gauss cherchait la valeur de la quantité inconnue X qui rendrait maximale la probabilité d'observation de mesures observées x_1 , x_2 , … , x_n et donc des erreurs correspondantes.

observations, il utilise la méthode des moindres carrés qui permet d'obtenir la moyenne arithmétique des observations pour valeur la plus vraisemblable, et la courbe « en cloche » comme courbe des erreurs autour de cette valeur théorique (il obtient la loi normale en tant que solution d'une équation différentielle). Cette étude statistique n'est finalement que le seul lien entre Gauss et cette loi.

La synthèse entre ces deux approches imposera, dès 1810, la loi normale comme loi quasi universelle : en effet, même si la distribution individuelle des erreurs ne suit pas une loi normale, celle des moyennes des erreurs suit sous certaines conditions (indépendance, lois identiques), une loi normale.

IV. Pourquoi « normale » ?

Puisqu'il y a plusieurs auteurs considérés comme pères de la loi normale, les noms « loi gaussienne » et « loi de Gauss » sont utilisés dans la littérature allemande et anglo-saxonne, alors que le nom « loi de Laplace » est utilisé par les mathématiciens français. Les noms « loi de Laplace-Gauss » et « loi de Gauss-Laplace » sont également parfois utilisés.

De nombreux auteurs ont donné des noms différents à cette loi : Adolphe Quetelet l'appelait « courbe des possibilités » ou « loi des erreurs accidentelles » ; Bartel Leendert van der Waerden (1967) l'appelait « courbe de Quetelet » ; Francis Galton (1877) parlait de « loi de fréquence des erreurs », de « loi de déviation d'après une moyenne », ou encore d'une courbe « de forme parfaitement normale » : c'est la première apparition du terme « normal » en tant qu'adjectif.

En 1893, ***Karl Pearson*** donne le nom de loi normale à la loi appelée par l'école française *deuxième loi de Laplace* ou *loi de Laplace-Gauss* ou *loi de Gauss* par l'école anglo-saxonne… Pearson explique donc le choix du terme « normal » pour la loi et la courbe par la facilité de ne pas fixer de paternité. Il reconnaîtra en 1920 que ce nom était inadéquat et qu'il « a le désavantage de conduire les gens à croire que toutes les autres distributions de fréquences sont en un sens ou un autre anormales ».

Autour de 1950, la commission de terminologie statistique de l'Afnor (Association Française de NORmalisation), probablement emmené par Fréchet, décida de normaliser le terme « loi de Laplace ».
Le polytechnicien André Pallez ajoute : « la Commission, considérant que Laplace a découvert la loi qui devrait porter son nom et qui porte celui de Gauss, à une époque où Gauss était encore un jeune enfant, a rétabli la vérité en rendant à Laplace l'hommage qui lui était dû. ».

Cependant, aujourd'hui au 21^{e} siècle, les deux noms les plus utilisés sont « loi de Gauss » et « loi normale ». Le nom de Gauss est resté plus que les autres grâce, entre autres, à l'influence de l'ouvrage pédagogique de Joseph Bertrand publié à la fin du 19^{e} siècle.

V. Quelques infos sur les mathématiciens contributeurs

Sources :
– « La courbe de Gauss ou le théorème de Bernoulli raconté aux enfants », Bernard Bru (http://msh.revues.org/3556)
– « La courbe de Gauss. D'où vient-elle ? A quoi sert-elle ? », Brigitte Chaput (http://apmep.poitiers.free.fr/IMG/pdf/LA_COURBE_DE_GAUSS-2.pdf)
– « Legendre et la méthode des moindres carrés » (http://www.bibnum.education.fr/files/legendre-analyse.pdf)
– « Plaidoyer pour la loi normale », Aimé Fuchs (http://www-irma.u-strasbg.fr/~foata/fuchs/FuchsNormale.pdf)
– Wikipédia et sites divers sur internet, dont http://www.bibmath.net/ et http://www-history.mcs.st-and.ac.uk/BiogIndex.html

DE MOIVRE

Abraham de Moivre (1667-1754) est né à Vitry-le-François. Fils de chirurgien, il bénéficie d'une bonne éducation qui le conduit vers les sciences. Mais il est protestant, et la révocation de l'Édit de Nantes, en 1685, l'oblige à s'installer à Londres avec son frère. Il y vivra la dure condition d'immigré : étant étranger, il ne parviendra jamais à obtenir de poste à l'Université, et devra gagner sa vie en donnant des cours particuliers, ou en utilisant son esprit vif et brillant pour résoudre toute sorte de problèmes dans les pubs.

L'apport de De Moivre est fondamental en probabilités, et son ouvrage, *Doctrine of chance*, paru en 1718, est la plus importante publication dans ce domaine entre les travaux de Pascal et Fermat, vers 1650, et ceux de Laplace, 50 ans après de Moivre. Ainsi, c'est dans cet ouvrage que de Moivre explique comment calculer la probabilité d'un événement aléatoire qui dépend de plusieurs autres événements : c'est la formule des probabilités composées[3]. De Moivre est aussi le premier à s'intéresser à la convergence des variables aléatoires, sous l'optique suivante : dans quelle mesure peut-on être sûr que lorsque l'on lance un grand nombre de fois un dé, la fréquence observée d'apparition du nombre 6 tend vers la probabilité théorique ? C'est une question essentielle à résoudre pour les problèmes de modélisation. De Moivre montre en particulier que la loi binomiale tend, en un certain sens, vers la loi normale, la fameuse loi "à la courbe en cloche".

De Moivre s'intéresse aussi aux applications pratiques des probabilités et statistiques. Il dresse ainsi des tables de mortalité précises, et donne des formules qui permettent de calculer équitablement le montant d'une rente viagère.

Une jolie légende entoure la mort de de Moivre, survenue le 27 novembre 1754 à Londres, dans la pauvreté. On raconte que de Moivre s'était rendu compte qu'il dormait chaque nuit un quart d'heure supplémentaire. S'aidant de cette suite arithmétique, il avait deviné le jour de sa mort, celui où il dormirait pendant 24 heures… et il ne s'était pas trompé !

3 par exemple $p(A\cap B)=p_B(A)\times p(B)$

LAPLACE

Normand et fils de cultivateur, Pierre-Simon Laplace (1749-1827) était l'un des principaux scientifiques de la période napoléonienne, apportant des contributions fondamentales dans différents champs des mathématiques, de l'astronomie et de la théorie des probabilités.

Laplace doit son éducation à l'intérêt de quelques riches voisins pour ses capacités et pour sa belle prestance. A 18 ans, il arrive à Paris avec une lettre de recommandation pour rencontrer le mathématicien d'Alembert, mais ce dernier refuse de rencontrer l'inconnue. Mais Laplace insiste, envoyant à d'Alembert un article qu'il a écrit sur la mécanique classique. D'Alembert en est si impressionné qu'il est tout heureux de patronner Laplace. Il lui obtient un poste d'enseignement en mathématiques. À la Révolution, il participa à l'organisation de l'École Normale et de l'École Polytechnique. Bonaparte lui confira le ministère de l'Intérieur, mais seulement pour 6 mois.

À la différence de beaucoup d'autres mathématiciens, Laplace ne donne pas aux mathématiques un statut particulier, il y voit plutôt un instrument utile pour la recherche scientifique et pour les problèmes pratiques. L'œuvre la plus importante de Laplace concerne le calcul des probabilités et la mécanique céleste.
On rapporte que, feuilletant la *Mécanique céleste* (ouvrage de Laplace, ayant pour objectif de résumer toutes les connaissances de l'époque sur la gravitation universelle) Napoléon fit remarquer à Laplace qu'il n'y était nulle part fait mention de Dieu. "Je n'ai pas eu besoin de cette hypothèse", rétorqua le savant.

GAUSS

Johann Carl Friedrich Gauss (1777-1855) était mathématicien, astronome et physicien allemand. Il a apporté de très importantes contributions à ces trois domaines. Surnommé « le prince des mathématiciens », il est considéré comme l'un des plus grands mathématiciens de tous les temps, et la qualité extraordinaire de ses travaux scientifiques était déjà reconnue par ses contemporains.
Il est à la fois le dernier des classiques, et le premier des modernes, c'est-à-dire qu'il a résolu les problèmes les plus classiques avec les méthodes les plus modernes. Par exemple, il démontra comment partager une tarte en 17 parts égales à l'aide d'une règle et d'un compas, ce qui était un problème ouvert depuis les Grecs. Mieux, il démontra pour quels nombres ce partage en parts égales est possible.

Gauss était un génie particulièrement précoce. On raconte (prudence !) qu'à un très jeune âge (moins de 10 ans), le maître demanda à ses élèves turbulents de calculer $1 + 2 + \ldots + 100$, pensant que cela les occuperait assez longtemps, et Gauss inscrivit rapidement le résultat sur son ardoise : ce n'est pas qu'il fut un génial calculateur, mais il avait trouvé une formule générale pour calculer de telles sommes, en remarquant que $1 + 100 = 2 + 99 = 3 + 98 = \ldots = 101$.

À l'université, à 19 ans, il fut le premier à démontrer la loi de réciprocité quadratique[4], ce que ni Euler, ni Legendre n'avaient réussi. Au cours de sa vie, il en donnera huit preuves !! Parmi ses autres prouesses, on peut citer la démonstration du théorème fondamental de l'algèbre *(tout polynôme non constant, à coefficients complexes, admet au moins une racine)* dans sa thèse de doctorat en 1799, l'invention de la théorie des congruences… Mais le génie de Gauss se manifesta dans d'autres domaines : on lui doit d'importants travaux en électricité, en optique, en théorie du potentiel. Ainsi, le "gauss" est devenu l'unité d'induction magnétique.

4 La loi de réciprocité quadratique établit des liens entre les nombres premiers (plus précisément, elle décrit la possibilité d'exprimer un nombre premier comme un carré, à un multiple près d'un autre nombre premier). Conjecturée par Euler et explicitée par Legendre, elle a été correctement démontrée pour la première fois par Gauss en 1801. Elle est considérée comme un des théorèmes les plus importants de la théorie des nombres, et a de nombreuses généralisations.

Il achève sa carrière de mathématicien en 1849, à l'occasion d'un jubilé en son honneur.
Peu à peu, sa santé se détériore, et il meurt à Göttingen le 23 février 1855 pendant son sommeil.

QUÉTELET

Adolphe Quételet (1796-1874), mathématicien, statisticien et astronome belge a été le premier à appliquer la distribution normale à des données sociales et biologiques. Très jeune, il fait d'intéressants travaux de géométrie, d'astronomie et de météorologie, 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).

En 1846, Quételet, après vingt ans de travaux statistiques divers, publie, à son tour, une version grand public des travaux statistiques du moment, enrichie d'exemples très remarquables empruntés aux statistiques du temps, qu'il connaît mieux que personne.

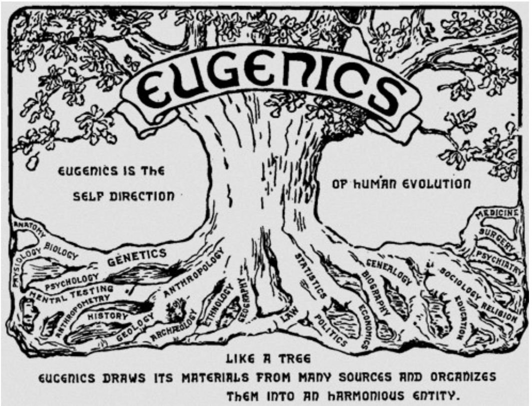
L'une des plus curieuses applications consiste à étudier un tableau, publié par l'armée britannique vers 1840, donnant les tours de poitrine de 5738 soldats écossais. Sans doute une commande de l'intendance militaire en vue de la confection de tenues aux bonnes tailles. Plutôt que se lancer dans de longs commentaires, Quételet a l'idée de comparer le tableau des tours de poitrine à celui de la loi normale correspondante.
Comment Quételet 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.

L'interprétation donnée par Quételet ne lui a pas survécu puisque l'on ne considère plus que la moyenne est l'idéal à atteindre par la nature et que les observations situées de part et d'autre de la moyenne représentent l'erreur, c'est-à-dire un écart par rapport à l'idéal de la nature.

GALTON

Issu d'une famille de scientifiques, Francis Galton (1822-1911) était un homme de science britannique.
Il fut anthropologue, explorateur, géographe, inventeur, météorologue, proto-généticien, psychométricien et statisticien.
Il fut l'un des pionniers en statistique, dans un but purement utilitaire. Ses travaux dans le domaine des statistiques restèrent cependant secondaires pour Galton, à côté de ses études sur l'origine des espèces : il voulait justifier la transmission des possibilités intellectuelles par l'hérédité pour améliorer l'espèce humaine. Il est par ailleurs considéré comme le fondateur de l'eugénisme*. Il a également mis en place de façon systématique la méthode d'identification des individus par empreintes digitales.

Il affirmera l'omniprésence de la loi normale dans la nature physique, comme biologique. C'est, par ailleurs, lui qui a donné à la distribution normale un rôle central dans la théorie sur les facultés mentales.



* L'eugénisme peut être défini comme l'ensemble des méthodes et pratiques visant à transformer le patrimoine génétique de l'espèce humaine, dans le but de le faire tendre vers un idéal déterminé. Il peut être le fruit d'une politique délibérément menée par un État. Il peut aussi être le résultat collectif d'une somme de décisions individuelles convergentes prises par les futurs parents, dans une société où primerait la recherche de l'« enfant parfait », ou du moins indemne de nombreuses affections graves.

Le terme *eugenics* a été employé pour la première fois en 1883 par le scientifique Galton dont les travaux ont grandement participé à la constitution et à la diffusion de l'idéologie eugéniste. Mené par des scientifiques et des médecins, le mouvement de promotion de l'eugénisme qui se met en place au tournant du 20^{e} siècle milite en faveur de politiques volontaristes d'éradication des caractères jugés handicapants ou de favorisation des caractères jugés bénéfiques. Son influence sur la législation s'est traduite principalement dans trois domaines : la mise en place de programmes de stérilisations contraintes, le durcissement de l'encadrement juridique du mariage et la restriction de l'immigration.

Dans la période contemporaine, les progrès du génie génétique et le développement des techniques de procréation médicalement assistée ont ouvert de nouvelles possibilités médicales (diagnostic prénatal, diagnostic préimplantatoire…) qui ont nourri les débats éthiques concernant la convergence des techniques bio-médicales et des pratiques sélectives.

PEARSON

Karl Pearson (1857-1936), mathématicien britannique, est un des fondateurs des statistiques modernes.
Il est aujourd'hui principalement connu pour avoir développé le test du Khi-deux[5], et pour avoir donné (en 1893) le nom de loi *normale* à la loi étudiée dans ce cours.

Fils d'un avocat en pleine ascension sociale, Pearson, devenu après de brillantes études professeur de mathématiques, est un pur produit de la « méritocratie », l'archétype de ces professions intellectuelles montantes, qui d'une part se dégagent des classes laborieuses et d'autre part se heurtent aux autorités traditionnelles. Socialiste élitiste et libre penseur positiviste[6], il laisse une œuvre scientifique qui ne peut se comprendre que dans le contexte du Royaume-Uni capitaliste et victorien.

C'est à l'âge de 33 ans que Pearson se tourne vers la statistique. Francis Galton vient de publier son ouvrage *Natural Inheritance* (L'héritage naturel) et Pearson, à sa suite, va appliquer les méthodes statistiques à l'étude de la sélection naturelle de Darwin dans le cadre des théories de l'eugénisme alors en vogue parmi les nouvelles élites.
À l'instar de Galton auquel le lie une indéfectible amitié, Pearson pense qu'il est possible et même éminemment souhaitable d'améliorer la race humaine en sélectionnant et favorisant les plus doués de ses représentants comme le fait la sélection naturelle pour les animaux. L'analyse statistique doit lui permettre

5 Test statistique permettant de tester l'adéquation d'une série de données à une famille de lois de probabilités ou de tester l'indépendance entre deux variables aléatoires.

6 Le terme positivisme désigne un ensemble de courants qui considère que seules l'analyse et la connaissance des faits vérifiés par l'expérience peuvent expliquer les phénomènes du monde. La certitude en est fournie exclusivement par l'expérience scientifique. Il rejette l'introspection, l'intuition et toute approche métaphysique pour expliquer la connaissance des phénomènes.

de mesurer la détermination héréditaire des caractéristiques physiques et psychiques de l’homme et leur amélioration.

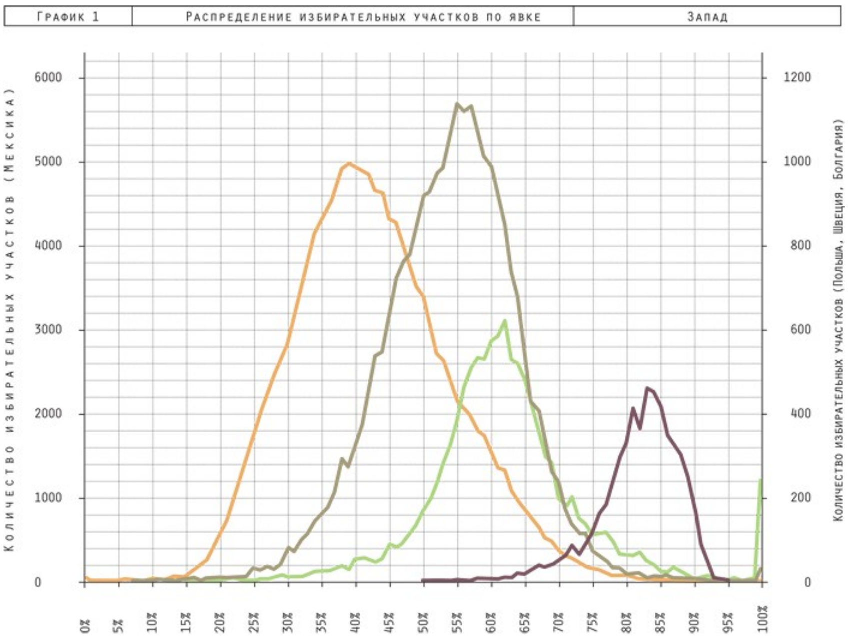
Karl Pearson s’emploiera toute sa vie à la promotion indissociablement liée pour lui de la statistique et de ses convictions eugénistes. Éditeur de revues, il publie de nombreuses tables statistiques et il forme au sein de son laboratoire de l’University College de Londres plusieurs générations de brillants statisticiens.

« Nous sommes une distribution normale »

« Pour la distribution normale »

« Nous ne croyons pas Churov[7] !
nous croyons Gauss ! »

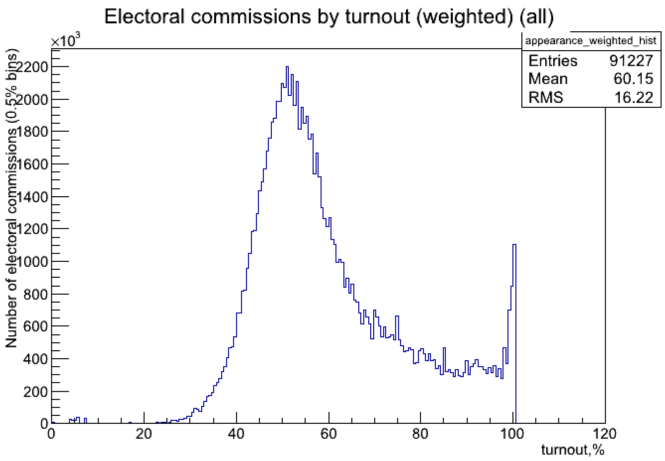
Le 4 décembre 2011, en Russie, les citoyens ont été appelés aux urnes afin d'élire leurs députés au cours d'un scrutin législatif. Plusieurs cas de fraudes en faveur de Russie unie sont alors dénoncés par l'opposition et des médias indépendants, des manifestations ont lieu dans les jours qui suivent.



Sur le graphique ci-dessus, on trouve en abscisse le pourcentage de votants, et en ordonnée le nombre de bureaux de vote pour quatre élections.

De gauche à droite : les législatives mexicaines de 2009 (courbe orange), le deuxième tour de la présidentielle polonaise en 2010 (en gris), les législatives bulgares en 2009 (en vert) et les législatives suédoises en 2010 (en mauve). Chaque pic reflète une tendance de participation au vote national (respectivement autour de 40 %, 55 %, 62 % et 82 %).

Ce graphique montre la participation par commissions électorales lors des élections législatives de 2011 en Russie.



Les circonscriptions ayant enregistré un taux de participation plus fort que la moyenne apparaissent bien plus nombreuses que celles dans lesquelles les électeurs se sont peu rendus aux urnes. On note un pic de bureaux de vote ayant enregistré un taux de participation proche des 100 %, à droite.

7 Vladimir Churov était le président de la « commission électorale centrale ».