IREM

des Pays de la Loire

LES MATHÉMATIQUES NE SE SONT PAS FAITES EN UN JOUR ...

Promenades historiques

Quatrième promenade : De l'urne de Bernoulli à l'Homme Moyen.

Septembre 2003

INSTITUT DE RECHERCHE SUR L'ENSEIGNEMENT DES MATHÉMATIQUES DES PAYS DE LA LOIRE

2, rue de la Houssinière • BP 92208
44322 NANTES CEDEX 03
Tel. 02 51 12 59 41
Site : www.irem.sciences.univ-nantes.fr/

Préambule

Nous vous proposons, au travers d'une série de petits cahiers, quelques promenades historiques au fil des mathématiques.

Nous nous adressons autant aux élèves qu'aux professeurs ; chacun, nous l'espérons, pourra y trouver son compte.

Pour cela, nous avons choisi un langage simple, et vous trouverez, au long des pages, des activités expérimentées par nos élèves. Ces activités permettent soit d'introduire une notion au programme, soit de la prolonger, soit aussi d'acquérir une culture mathématique.

Les mathématiques ne se sont pas faites en un jour ; elles continuent de se construire. Elles ont une histoire pleine de bonds et de rebonds. En lisant cette histoire, l'élève comprendra que ses difficultés furent peut-être aussi celles des plus grands mathématiciens. Il découvrira comment certaines notions ont évolué et pourquoi. Il fera des mathématiques autrement.

Nous appuierons ces promenades sur la lecture de textes originaux, en proposant, lorsque cela semble nécessaire pour la compréhension, des traductions en termes plus modernes.

Nous indiquerons les niveaux d'accessibilité, et les prérequis, le cas échéant. Mais nous avons pris le parti de ne privilégier aucun niveau, du collège au lycée, en passant par le lycée professionnel.

Nous vous souhaitons d'agréables promenades mathématiques ; le rêve est peut-être au bout du chemin.

Ce fascicule s'adresse plutôt à des élèves de lycée. Le premier chapitre est abordable par des élèves de premières

. Le deuxième chapitre est assez difficile et sera peut-être réservé aux terminales. Les chapitres II et IV sont en grande partie abordables dès la seconde ; le chapitre V peut s'adresser à tous les niveaux, même de collège.

Les auteures de ce fascicule, Marie-Céline Comairas et Anne Boyé, remercient Xavier Lefort pour sa collaboration active.

Jacques Bernoulli

1. De l'art de conjecturer

Si Pascal et Fermat ont donné naissance au calcul des probabilités, c'est Jacques Bernoulli (1654-1705) qui, dans l' *Ars conjectandi* publié en 1713, 8 ans après sa mort, par son neveu Nicolas Bernoulli, apporte la première grande contribution théorique à ce calcul. C'est lui qui fait le lien entre la probabilité *a priori* d'un événement, probabilité que l'on peut calculer **avant** la réalisation d'une expérience aléatoire (par exemple, la probabilité de tirer un as d'un jeu de 32 cartes est égale à $\frac{1}{8}$), et la probabilité *a posteriori*, c'est-à-dire la probabilité que l'on déduit de la fréquence d'apparition de cet événement **après** qu'on ait répété un grand nombre de fois l'expérience aléatoire.

Ce n'est pas seulement l'aspect mathématique du calcul des probabilités qui intéresse Jacques Bernoulli ; son objectif est beaucoup plus large : il veut montrer *« comment les bases de l'art de conjecturer peuvent s'appliquer à la vie civile, morale et politique. »* Malheureusement il ne réussira pas à l'atteindre complètement et son traité restera inachevé, ce qui explique le fait qu'il n'ait pas été édité de son vivant.

L'*Ars conjectandi* comporte quatre parties que Nicolas Bernoulli présente ainsi dans la préface :

> [...]L'Auteur a divisé cette Œuvre en quatre parties. La Ière contient le Traité de l'Illustre Huygens *De Ratiociniis in Aleae Ludo*, avec des notes, Traité que mon oncle a jugé devoir être mis en tête de son ouvrage, car on y trouve les premiers éléments de l'Art de Conjecturer. La IIème partie embrasse la Théorie des Permutations et des Combinaisons, théorie absolument nécessaire pour mesurer les probabilités ; dans la IIIème partie il en a développé les applications aux divers Tirages au sort et Jeux de Hasard. Quant à la IVème partie par laquelle il aurait voulu montrer l'application et le rapport des parties précédentes aux affaires civiles, morales et économiques, affligé depuis longtemps d'une mauvaise santé et enfin prévenu par la mort elle-même, il l'a laissée inachevée.

Cette quatrième partie, laissée inachevée donc, est cependant la plus célèbre car elle contient, entre autres, la modélisation du réel par une urne contenant des pierres blanches et noires et surtout l'énoncé de la loi des grands nombres.

Nous vous proposons quelques extraits de cette quatrième partie[1].

L'*Ars conjectandi* traite des probabilités, mais qu'est-ce qu'une probabilité ? Voici comment Jacques Bernoulli la définit :

> [...] Ce qui par révélation, raison, sensation, expérience, témoignage de nos propres yeux ou autrement est tellement évident que nous ne pouvons aucunement douter de son existence présente ou future jouit d'une certitude totale et absolue. Tout le reste acquiert dans nos esprits une mesure moins parfaite, plus ou moins grande, selon que les probabilités sont plus ou moins nombreuses qui nous persuadent qu'une chose est, sera ou a été.
>
> La *probabilité* est en effet un degré de la certitude et en diffère comme la partie diffère du tout. Evidemment, si la certitude intégrale et absolue, que nous désignons par la lettre a ou par l'unité 1, est constituée de - supposons par exemple - cinq probabilités ou parties, dont trois militent pour qu 'un événement existe ou se produise, les autres s'y opposant : nous dirons que cet événement a $\frac{3}{5}a$, ou $\frac{3}{5}$ de certitude.
>
> Sera donc dit *plus probable* qu'autre chose ce qui a une plus grande part de certitude, bien qu'en fait, d'après le langage courant, soit dit *probable*, ce dont la probabilité dépasse notablement la moitié de la certitude. Je dis : *notablement* ; car ce qui équivaut à la moitié de la certitude environ est dit *douteux* ou ambigu. Ainsi ce qui a $\frac{1}{5}$ de certitude est plus probable que ce qui n'en a qu' $\frac{1}{10}$; même si ni l'un ni l'autre n'est en fait probable.
>
> Ce à quoi est attribuée une part de certitude, si peu même que ce soit, cela est *possible* ; lorsque cette part est nulle ou infiniment petite cela est *impossible*. Ainsi est possible ce qui a $\frac{1}{20}$ ou $\frac{1}{30}$ de certitude.

On voit par ce texte la difficulté liée au vocabulaire utilisé : le mot *probable* n'a pas le même sens en mathématique et dans la vie courante. Par ailleurs la définition que donne Jacques Bernoulli de la probabilité est-elle vraiment claire et précise ?

C'est au début du chapitre II que Jacques Bernoulli expose l'objet de cette quatrième partie :

[1] L'*Ars conjectandi* est écrit en latin. La traduction utilisée ici est celle de Norbert Meusnier éditée par l'IREM de Rouen.

Ce qui est certain et indubitable, nous disons que nous le *savons* ou que nous le *comprenons* ; tout le reste nous le *conjecturons* ou le *présumons*.

Conjecturer quelque chose, c'est mesurer sa probabilité : ainsi *l'art de conjecturer* ou la *stochastique* se définit pour nous comme l'art de mesurer aussi exactement que possible les probabilités des choses. Le but est que dans nos jugements et dans nos actions nous puissions toujours choisir ou suivre le parti que nous aurons découvert comme meilleur, préférable, plus sûr ou mieux réfléchi. C'est en cela seulement que réside toute la sagesse du Philosophe et toute la sagacité du Politique.

Les probabilités sont estimées d'après le *nombre* et aussi le *poids des arguments* qui de quelque manière prouvent ou révèlent que quelque chose est, sera ou a été. En outre, par le *poids* j'entends la force de ce qui prouve.

Dans le chapitre IV il indique la difficulté de déterminer les probabilités :

[...] On en est ainsi venu à ce point que pour former selon les règles des conjectures sur n'importe quelle chose il est seulement requis d'une part que les nombres de cas soient soigneusement déterminés, et d'autre part que soit défini combien les uns peuvent arriver plus facilement que les autres. Mais c'est ici enfin que surgit une difficulté, nous semble-t-il : cela peut se voir à peine dans quelques très rares cas et ne se produit presque pas en dehors des jeux de hasard que leurs premiers inventeurs ont pris soin d'organiser en vue de se ménager l'équité, de telle sorte que fussent assurés et connus les nombres de cas qui doivent entraîner le gain ou la perte, et de telle sorte que tous ces cas puissent arriver avec une égale facilité. En effet lorsqu'il s'agit de tous les autres résultats, dépendant pour la plupart soit de l'œuvre de nature soit de l'arbitre des hommes, cela n'a pas du tout lieu. Ainsi, par exemple, les nombres de cas sont connus lorsqu'il s'agit des dés, car pour chacun des dés les cas sont manifestement aussi nombreux que les bases, et ils sont tous également enclins à échoir ; car à cause de la similitude des bases et du poids uniforme des dés il n'y a point de raison pour qu'une des bases soit plus encline à échoir que l'autre, comme cela arriverait si les bases étaient de formes dissemblables, ou si le dé était constitué d'un côté d'une matière plus lourde que de l'autre. [...] Mais qui donc parmi les mortels définira par exemple le nombre de maladies, qui sont autant de cas, qui ont le pouvoir d'envahir les innombrables parties du corps humain à l'âge qu'on voudra, et qui ont le pouvoir de nous apporter la mort ? Qui définira combien est plus facile à celle-ci qu'à celle-là, la peste ou l'hydropisie, l'hydropisie ou la fièvre, d'anéantir un homme, en sorte qu'à partir de là puisse être formée une conjecture sur un état futur de vie ou de mort ?

[...] Mais à la vérité ici s'offre à nous un autre chemin pour obtenir ce que nous cherchons. Ce qu'il n'est pas donné d'obtenir *a priori* l'est du moins *a posteriori*, c'est-à-dire qu'il sera possible de l'extraire en observant l'issue de nombreux exemples semblables ; car on doit présumer que, par la suite, chaque fait peut arriver et ne pas arriver dans le même nombre de cas qu'il avait été constaté auparavant, dans un état de choses semblables, qu'il arrivait ou n'arrivait pas.

[...] Enfin il ne peut échapper à personne que, pour juger par ce moyen de quelque événement, il ne suffirait pas d'avoir fait choix d'une ou deux expériences, mais qu'il serait requis une grande quantité d'expériences.[...]

Et un peu plus loin il pose le problème dont la solution sera donnée dans le chapitre suivant sous la forme d'un théorème, que Poisson nommera en 1857 la « loi des grands nombres » :

> [...] Il reste assurément à chercher si, en augmentant ainsi le nombre des observations, nous augmentons continuellement la probabilité d'atteindre le rapport réel entre les nombres de cas qui font qu'un événement peut arriver et le nombre de ceux qui font qu'il ne peut arriver, de sorte que cette probabilité dépasse enfin un degré quelconque donné de certitude ; ou si le Problème , pour ainsi dire, a son Asymptote, c'est-à-dire s'il existe un degré de certitude qu'il n'est jamais possible de dépasser, de quelque manière qu'on multiplie les observations, d'avoir découvert le vrai rapport des cas ; que, par exemple, nous ne pouvons jamais obtenir de certitude au-delà de la moitié, ou de $\frac{2}{3}$ ou de $\frac{3}{4}$. Un exemple rendra clair ce que je voudrais dire.

C'est alors qu'il propose d'éclairer son propos en prenant pour modèle une urne[2] :

> Je suppose que, dans une urne, à ton insu soient placées trois mille pierres blanches et deux mille pierres noires ; je suppose que pour connaître leurs nombres par expérience tu tires une pierre après l'autre (en replaçant cependant chaque fois la pierre que tu as tirée avant de choisir la suivante, pour que le nombre des pierres ne diminue pas dans l'urne) ; tu observes combien de fois sort une pierre blanche et combien de fois une noire. On demande si tu peux le faire tant de fois qu'il devienne dix fois, cent fois, mille fois, etc.. ; plus probable (c'est-à-dire qu'il devienne moralement certain) que le nombre de fois où tu choisis une pierre blanche et le nombre de fois où tu choisis une pierre noire soient dans ce même rapport sesquialtère[3] où se complaisent à être entre eux les nombres de pierres ou de cas, plutôt que dans tout autre rapport différent de celui-ci. Car si cela ne se produisait pas, j'avoue que c'en serait fait de notre effort pour rechercher expérimentalement le nombre de cas. Mais si nous l'obtenons et si nous acquérons enfin par ce moyen la certitude morale (et je montrerai dans le chapitre suivant que cela aussi se produit réellement), nous aurons trouvé *a posteriori* les nombres de cas presque comme s'ils nous étaient connus a priori.

[2] Pour J.Bernoulli le modèle de l'urne est fondamental : c'est à la fois un dé généralisé qui garantit l'égalité des cas et un nombre de cas aussi grand qu'on veut et le modèle de toutes les situations imaginables. Il permet de modéliser les situations les plus simples comme les plus complexes.

[3] se dit de deux quantités dont l'une contient l'autre une fois et demie.

Le problème présenté plus haut est précisé grâce au modèle de l'urne :

Mais pour que cela ne soit pas compris autrement qu'il ne convient, il faut bien noter ce qui suit ; je voudrais que le rapport entre les nombres de cas, que nous entreprenons de déterminer expérimentalement, ne fût pas pris de façon nette et sans partage (car ainsi c'est tout le contraire qui arriverait et il deviendrait d'autant moins probable de découvrir le vrai rapport qu'on ferait de plus nombreuses observations), mais je voudrais que le rapport fût admis avec une certaine latitude, c'est-à-dire compris entre une paire de limites, pouvant être prises aussi rapprochées qu'on voudra. Assurément, si dans l'exemple des pierres proposé plus haut nous prenons les deux rapports $\frac{301}{200}$ et $\frac{299}{200}$, ou $\frac{3001}{2000}$ et $\frac{2999}{2000}$,etc.. dont le sesquialtère est très près et du plus grand et du plus petit, on montrera que l'on peut arriver à ce que le rapport trouvé grâce à des expériences recommencées de nombreuses fois tombe entre ces limites du rapport sesquialtère plus probablement, de toute probabilité donnée, qu'en dehors.

Enfin, dans le chapitreV, intitulé « Solution au problème précédent », Jacques Bernoulli expose et démontre le théorème suivant :

[...] Soit donc le nombre de cas fertiles au nombre de cas stériles précisément ou approximativement dans le rapport $\frac{r}{s}$ et qu'il soit en conséquence, au nombre de tous dans le rapport $\frac{r}{r+s}$ ou $\frac{r}{t}$, rapport qu'encadrent les limites $\frac{r+1}{t}$ et $\frac{r-1}{t}$. Il faut montrer que l'on peut concevoir des expériences en un nombre tel qu'il soit plus vraisemblable d'autant de fois que l'on veut (soit c) que le nombre des observations tombe à l'intérieur de ces limites plutôt qu'en dehors, c'est-à-dire que le nombre des observations fertiles soit au nombre de toutes les observations dans un rapport ni plus grand que $\frac{r+1}{t}$, ni plus petit que $\frac{r-1}{t}$.

Voici un énoncé plus actuel de ce théorème :

Pour tout nombre ε strictement positif, la probabilité que, dans une succession de n épreuves de Bernoulli indépendantes de même probabilité de succès p, la fréquence Z_n des succès s'écarte de plus de ε de la probabilité p tend vers zéro lorsque n tend vers l'infini.

Plus précisément : $P(|Z_n - p| \geq \varepsilon) \leq \frac{pq}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}$

Cours de probabilités et de statistiques.C LEBOEUF , J-L ROQUE, J.GUEGAND. Ellipses.

Et enfin une version plus « triviale » :

> Il est très peu probable que, si l'on fait un nombre suffisamment grand d'expériences, la fréquence d'apparition d'un événement s'écarte notablement de sa probabilité.

> <u>Exercice</u> : Buffon, au cours d'une de ses expériences, a lancé 4040 fois une pièce de monnaie et il a obtenu 2049 fois pile. Déterminer un intervalle tel que la probabilité d'obtenir pile appartienne à cet intervalle avec une probabilité au moins égale à 0,95. La pièce de Buffon était-elle équilibrée ?

2 - De la facilité des erreurs

Ce chapitre est né de la lecture d'un ouvrage de Joseph Louis Lagrange, *Mémoire sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations, dans lequel on examine les avantages de cette méthode par le calcul des probabilités, et où l'on résout différents problèmes relatifs à cette matière*, ouvrage publié en 1774[1].

Il s'agit donc là encore plus de probabilité que de statistiques, mais le problème que se propose de résoudre Lagrange, relatif à l'évaluation des erreurs liées au choix de la moyenne entre des mesures, est un de ceux qui ont fondé la pensée statistique.

Lagrange ici exerce son talent de calculateur, en utilisant avec virtuosité les développements de ce que nous appelons le binôme de Newton. Il ne faudra pas se laisser rebuter par cet aspect du travail.

Joseph Louis Lagrange

[1] In Miscellanea Taurinensia, t V, 1770-1773.

Voici le problème que pose Lagrange au début de son mémoire :

> **PROBLEME :**
>
> "Quand on fait plusieurs observations d'un même phénomène dont les résultats ne sont pas tout à fait d'accord, on est sûr que ces observations sont toutes, ou au moins en partie, peu exactes, de quelque source que l'erreur puisse provenir ; alors on a coutume de prendre le milieu[2] entre tous ces résultats, parce que de cette manière, les différentes erreurs se répartissant également dans toutes les observations, l'erreur qui peut se trouver dans le résultat moyen devient aussi moyenne entre toutes les erreurs. Or, quoique tout le monde reconnaisse l'utilité de cette pratique pour diminuer, autant qu'il est possible, l'incertitude qui naît de l'imperfection des instruments et des erreurs inévitables des observations, j'ai cru cependant qu'il serait bon d'examiner et d'apprécier par le calcul les avantages qu'on peut espérer de retirer d'une semblable méthode ; c'est l'objet que je me suis proposé dans ce mémoire."

Pour aborder ce problème, Lagrange précise les conditions dans lesquelles il va conduire ses calculs, suivant "*la règle ordinaire du calcul des probabilités*", du rapport du nombre de cas favorables au nombre de tous les cas possibles. D'une certaine façon il va d'abord évaluer des probabilités "a priori", pour reprendre le vocabulaire de J. Bernoulli, supposant que les erreurs qui peuvent se glisser dans chaque observation sont données, que le nombre de cas qui donne ces erreurs est connu, ce qu'il nomme la "facilité" de chaque erreur, nom qui sera repris par Legendre et Gauss. Il suppose aussi que l'on connaît les limites entre lesquelles se trouvent toutes les erreurs et, à partir de la "loi de facilité" des erreurs (ce que nous nommerions loi de probabilité), il va essayer d'évaluer la probabilité que l'erreur commise en prenant le résultat moyen soit comprise entre des limites données.

A partir de là, ayant construit en quelque sorte un modèle, il montrera comment déterminer "a posteriori" la loi de probabilité des erreurs.

---

[2] Nous aurions tendance, implicitement, à remplacer le mot "milieu" par le mot "moyenne". Nous choisissons de conserver "milieu", qui ne posera pas de problème de compréhension pour la suite, et qui garde son côté intuitif, voire un peu flou, jusqu'à ce que les mathématiciens aient défini explicitement ce qu'ils entendent par moyenne. Contrairement à ce que l'on penserait, ce n'est pas si évident. (Voir Chapitre III)

Lagrange souligne que ce qui peut poser problème dans son travail est la complication de certains calculs, "dont on ne peut venir à bout que par des artifices particuliers".
C'est la raison pour laquelle nous nous garderons de le suivre pas à pas. Nous vous proposons de comprendre sa démarche, à partir de quelques exemples simples. Chacun sentira sans doute la richesse de la réflexion qui s'offre à nous.

Vous faites une observation dont le résultat "exact" serait ρ, que vous ne connaissez pas, bien sûr. Pour les calculs, Lagrange propose d'abord de supposer que l'on puisse faire une erreur d'une unité en plus ou en moins. Chaque observation pourra donc donner ρ, ou $\rho - 1$, ou $\rho + 1$.

Question 1 : On fait deux observations.

a) Si les deux résultats observés sont ρ et $\rho - 1$, quel est le "milieu" de ces deux résultats ? Quelle est l'erreur commise en prenant ce milieu ?
b) Faire un schéma représentant, lors de deux observations, tous les assemblages de deux résultats possibles, ainsi que les milieux correspondants à chacun de ces assemblages et l'erreur commise en prenant ces milieux.

Question 2 : On fait trois observations.

Faire de même un schéma représentant tous les assemblages de trois résultats possibles, les milieux correspondant à chacun de ces assemblages, et l'erreur commise en prenant ces milieux.

Nous proposons maintenant d'évaluer la probabilité qu'en prenant le milieu entre les résultats l'erreur ne surpasse pas un certaine fraction $\frac{m}{n}$ $(m < n)$, en valeur absolue.

Question 3 :

a) En utilisant le premier tableau, supposant que tous les assemblages de mesures sont équiprobables, évaluer la probabilité qu'en prenant le milieu entre deux observations, l'erreur ne surpasse pas $\frac{1}{2}$; puis $\frac{1}{3}$, en valeur absolue. (C'est-à-dire soit inférieure ou égale.)

b) En utilisant le deuxième tableau, répondre à la même question pour trois observations.

c) Comparant ces deux résultats, le choix du milieu est-il meilleur dans le premier cas ou dans le deuxième cas ? Est-ce la conclusion que vous attendiez ?

Voici ce que nous pouvons lire dans le mémoire de Lagrange :

> Ainsi, quoique la probabilité que l'erreur soit nulle puisse être plus petite lorsqu'on prend le résultat moyen de plusieurs observations que lorsqu'on prend le résultat de chaque observation en particulier, cependant, si l'on cherche que l'erreur ne surpasse pas $\frac{1}{2}$ ou $\frac{1}{3}$, ..., on trouvera que cette probabilité sera plus grande dans le premier cas que dans le second. En effet, dans le premier cas, il n'y a d'autres cas favorables que ceux où l'erreur est absolument nulle ; mais, dans le second, les cas favorables seront non seulement ceux où l'erreur est nulle, mais aussi ceux où l'erreur est $\frac{1}{2}$ ou $\frac{1}{3}$,... ; et c'est par cette considération qu'il est toujours plus avantageux de prendre le milieu entre les résultats de plusieurs observations que de s'en tenir au résultat de chaque observation en particulier.

Il s'agit dès lors d'évaluer, lors de n observations, la probabilité qu'en prenant le milieu des observations, l'erreur commise ne dépasse pas le rapport $\frac{m}{n}$, et de savoir si, lorsque le nombre n d'observations augmente, l'erreur commise en prenant le milieu diminue. Dans le cas le plus général, les erreurs ne seront pas équiprobables.

<u>Question 4</u> : On fait n observations, dans les mêmes conditions que précédemment, ρ étant la valeur "vraie" et les erreurs possibles étant de +1 ou -1.

Justifier que toutes les erreurs commises en prenant les milieux des observations sont éléments de $\{0, \pm\frac{1}{n}, \pm\frac{2}{n}, \pm\frac{3}{n}, ...1\}$.

Pour déterminer les probabilités que l'erreur commise en prenant le milieu ne dépasse pas $\frac{m}{n}$, Lagrange propose donc d'évaluer les probabilités de commettre une erreur de 0, $\pm\frac{1}{n}, \pm\frac{2}{n}, \ldots \pm\frac{m}{n}$, c'est-à-dire, d'une manière générale de $\pm\frac{\mu}{n}$.

Pour évaluer cette probabilité, il propose de raisonner sur un modèle : le lancer de n dés. Se plaçant dans le cas le plus général de la non équiprobabilité des erreurs, chaque dé est conçu pour rendre compte de la "facilité" de l'erreur 0, de l'erreur 1, de l'erreur -1.

Les n expériences ont lieu dans les mêmes conditions, donc les dés sont identiques. Chacun aura "a" faces marquées 0, "b" faces marquées 1 et "b" faces marquées - 1.

Il est en général admis, en effet, qu'il y a symétrie dans les erreurs commises lors de la mesure d'une grandeur[3].

Ainsi, chacun des dés aura a + 2b faces.

Lagrange, par ce modèle, se place donc dans un cas d'équiprobabilité, comme il l'avait annoncé au départ.

<u>Question 5</u> : Dans ces conditions, lors du lancer de n dés, quel est l'univers ? Quel est le nombre de résultats possibles, c'est-à-dire le nombre d'éventualités de l'univers ?

Pour évaluer la probabilité que l'erreur commise en prenant le milieu des n observations soit $\pm\frac{\mu}{n}$, il faut évaluer la probabilité d'amener + μ ou - μ points avec ces n dés.

<u>Question 6</u> : Etude d'un exemple.

Nous supposons que a = 1 et b = 3 ; c'est-à-dire que chaque dé a une face marquée 0, 3 faces marquées 1 et 3 faces marquées -1.

---

[3] Nous retrouverons cette idée de façon plus précise lors de l'examen de la loi des erreurs énoncée par Gauss. Ce principe est admis depuis au moins Galilée.

a) On lance 3 dés, c'est -à -dire que n = 3, ce qui revient à dire que l'on fait 3 observations.

Evaluer le nombre de façons d'obtenir ± 3, ± 2, ± 1, 0 points, puis les probabilités de chacun de ces événements.

Quelle est la probabilité que l'erreur commise (en valeur absolue), en prenant le milieu des valeurs observées, soit égale à 0 ? Inférieure ou égale à $\frac{1}{3}$? Inférieure ou égale à $\frac{2}{3}$?

b) On lance 4 dés.

Evaluer le nombre de façons d'obtenir ± 4, ± 3, ± 2, ± 1, 0 points, puis les probabilités de chacun de ces événements.

Quelle est la probabilité que l'erreur commise (en valeur absolue) en prenant le milieu des valeurs observées, soit égale à 0 ? inférieure ou égale à $\frac{1}{4}$? à $\frac{1}{2}$? à $\frac{3}{4}$?

c) On lance 5 dés

Répondre aux mêmes questions en transposant à n = 5.

d) Répertorier les résultats obtenus dans le tableau suivant :

Probabilité que l'erreur commise en prenant le milieu des observations ne surpasse pas, en valeur absolue :

Valeurs du nombre n des observations	0	$\frac{1}{n}$	$\frac{2}{n}$	$\frac{3}{n}$	$\frac{4}{n}$
3				....	....
4					....
5					

e) Comparer pour chacune des valeurs de n étudiées, les probabilités que l'erreur ne dépasse pas, en valeur absolue, $\frac{1}{3}$.

Vous venez d'effectuer des calculs quelque peu délicats, reconnaissons-le. Il s'agissait de pénétrer un peu dans le style de pensée de notre auteur. Lagrange lui-même pour commencer son raisonnement s'est rôdé sur des cas un peu plus simples. Voici par exemple le tableau qu'il a établit en choisissant a = b = 1.

VALEURS du nombre n des observations.	PROBABILITÉS QUE L'ERREUR NE SURPASSERA PAS LES FRACTIONS					
	$\pm\frac{0}{n}$.	$\pm\frac{1}{n}$.	$\pm\frac{2}{n}$.	$\pm\frac{3}{n}$.	$\pm\frac{4}{n}$.	$\pm\frac{5}{n}$.
1	$\frac{1}{3}$	...	...	...	...	...
2	$\frac{3}{9}$	$\frac{7}{9}$	...	...	...	...
3	$\frac{7}{27}$	$\frac{19}{27}$	$\frac{25}{27}$	...	...	...
4	$\frac{19}{81}$	$\frac{51}{81}$	$\frac{71}{81}$	$\frac{79}{81}$	...	...
5	$\frac{51}{243}$	$\frac{141}{243}$	$\frac{201}{243}$	$\frac{231}{243}$	$\frac{241}{243}$	...
6	$\frac{141}{729}$	$\frac{393}{729}$	$\frac{573}{729}$	$\frac{673}{729}$	$\frac{715}{729}$	$\frac{727}{729}$
....	...	...	...	...	...	...

A partir de ce tableau, revenant sur le fait que si l'on étudie la probabilité que l'erreur soit nulle il n'est pas forcément plus avantageux de faire plusieurs observations, il va s'attacher aux avantages de prendre le milieu entre plusieurs observations, et examiner si le nombre d'observations intervient.

"Pour rendre la chose plus sensible, nous allons rechercher les probabilités que l'erreur ne surpassera pas la fraction $\frac{1}{2}$. (...) On voit par là que la probabilité que l'erreur ne surpassera pas $\frac{1}{2}$ va en augmentant, à mesure que l'on prend un plus grand nombre d'observations, mais avec cette différence que la probabilité est plus grande pour deux observations que pour trois, pour quatre que pour cinq, et en général pour un nombre pair quelconque que pour le nombre impair qui le suit immédiatement ; de sorte que, dans l'hypothèse dont il s'agit, il est plus avantageux de ne prendre le milieu qu'entre un nombre quelconque pair d'observations."

Faites-vous les mêmes remarques à partir du tableau que vous avez réalisé ? Quelle hypothèse faite par Lagrange induit peut-être ce résultat ?

Lagrange fait une étude très générale, pour n quelconque, a et b quelconques. Pour évaluer la probabilité d'obtenir $\pm\ \mu$ points lors de n lancers, il utilise le développement de $\left[a+b(x+x^{-1})\right]^n$ avec la formule du binôme de Newton. Le nombre de façons d'obtenir μ est le coefficient de x^{μ} et le nombre de façons d'obtenir $-\mu$ est le coefficient de $x^{-\mu}$.

C'est une idée inattendue mais très astucieuse[4]. En effet pour obtenir μ, il faut que la répartition entre 0, 1 et -1 donne un "dépassement" de μ. Dans le développement de $(x + x^{-1})^n$ tous les termes en $(x \times x^{-1})^{\alpha}$ vont correspondre à ceux où il y a autant de faces 1 que de faces – 1. Ceux en $x^{\alpha}x^{-\beta}$ donnent $x^{\alpha-\beta}$, et si l'on considère $\alpha - \beta = \mu$ ou $\alpha - \beta = -\mu$ selon que ce résultat soit positif ou négatif, nous aurons le "dépassement" en 1 ou en – 1 .

<u>Question 7 :</u> vérifier ce résultat sur les cas étudiés.

Le nombre de cas total, le cardinal de l'univers, est $(a + 2b)^n$; donc on obtient la probabilité que l'erreur soit $\pm\dfrac{\mu}{n}$ en divisant le coefficient précédent par $(a + 2b)^n$.

Ceci étant, ce calcul, sur une expression générale, donc littérale, est très ardu. Lagrange le mène à bien par une succession d'astuces, et l'on sent ici le calculateur génial. Nous vous épargnerons ces calculs.

Il nous fait cependant remarquer que si l'on supposait que la probabilité de l'erreur 1 n'était pas exactement la même que l'erreur - 1 (donc s'il n'y avait pas symétrie des erreurs), le problème se résoudrait avec *la même facilité* en développant $(a+bx+cx^{-1})^n$ au lieu de $(a+b(x+x^{-1}))^n$.

Il généralise aussi au cas où l'erreur possible par rapport à ρ n'est pas forcément 0, 1, - 1 mais par exemple 0, r, - 1 ...

Lagrange s'intéresse maintenant à la probabilité que l'erreur moyenne ne dépasse pas, en valeur absolue, $\frac{\mu}{n}$.

Il nous prévient : "*Au reste, comme ce calcul est un peu long, nous nous contenterons de l'indiquer ici, pour mettre sur la voie ceux qui voudront pousser cette théorie plus loin.*"

Il poursuit alors le programme qu'il s'est tracé;

> "Problème V :
>
> 16 On suppose que chaque observation soit sujette à des erreurs quelconques données, et qu'on connaisse en même temps le nombre des cas où chaque erreur peut avoir lieu ; on demande la correction qu'il faudra faire au résultat moyen de plusieurs observations."

Comme Lagrange, appelons p, q, r, s ... les erreurs auxquelles chaque observation est sujette, et a, b, c, d ...les cas qui peuvent donner ces erreurs ; (a est le nombre de cas qui donnent l'erreur p, b l'erreur q, ...). Raisonnant comme précédemment, la probabilité que l'erreur du résultat moyen, au cours de n observations, soit $\frac{\mu}{n}$ sera : $\frac{M}{(a+b+c+...)^n}$ où M est le coefficient de x^μ dans le développement de $((ax^p + bx^q + cx^r + ...)^n$.

Utilisant les calculs fait précédemment, Lagrange établit que, dans ces conditions, l'erreur moyenne, qui est celle pour laquelle la probabilité est la plus grande est :

$$\frac{ap+bq+cr+...}{a+b+c+...}$$

C'est la correction qu'il faudra faire au résultat moyen de plusieurs observations.

<u>Question 8</u> : comment appelleriez-vous le nombre $\frac{ap+bq+cr+...}{a+b+c+...}$?

---

[4] Nous trouvons ici une de ces astuces de calcul que Lagrange nous a annoncées. Il a largement utilisé tous ces développements dans sa théorie des dérivées par exemple. Il est donc assez naturel qu'il y pense.

Question 9 : Voici le corollaire énoncé par Lagrange :

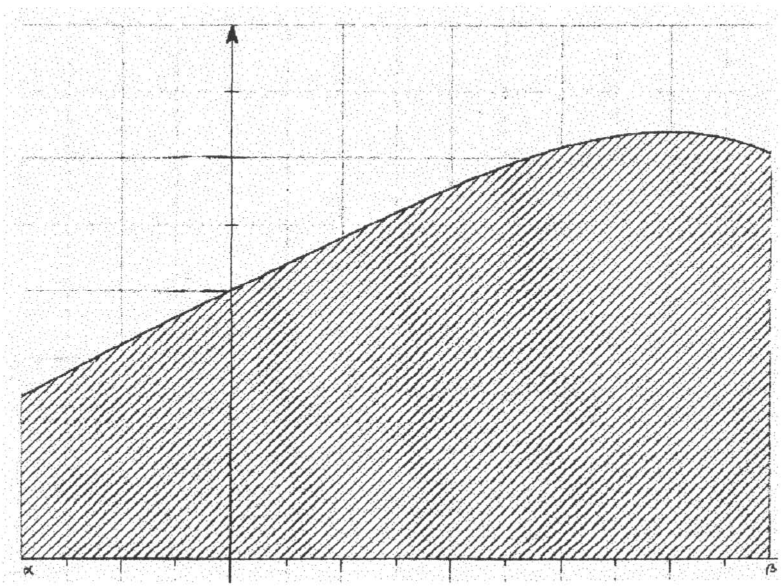
"17 Si l'on regarde les quantités a, b, c, ... comme des points appliqués à une droite indéfinie, à des distances égales à p, q, r, ... d'un point fixe pris dans cette droite, et qu'on cherche le centre de gravité de ces poids, la distance de ce centre au point fixe sera la correction qu'il faudra faire au résultat moyen de plusieurs observations."

Comment interprétez-vous ce corollaire ? Avez-vous envie d'utiliser d'autres mots ? Recherchez éventuellement quand sont apparus ces mots que vous voudriez utiliser, et les obstacles conceptuels auxquels Lagrange se heurte peut-être.

Vers les "lois continues" [5]:

Lagrange imagine alors que chaque observation puisse être sujette à toutes les erreurs entre deux bornes données ; par facilité nous noterons l'intervalle des erreurs possibles [α ; β] ; il imagine aussi que l'on connaisse "la courbe de facilité des erreurs"[6].

"Les abscisses étant supposées représenter les erreurs, les ordonnées représentent les facilités de ces erreurs."

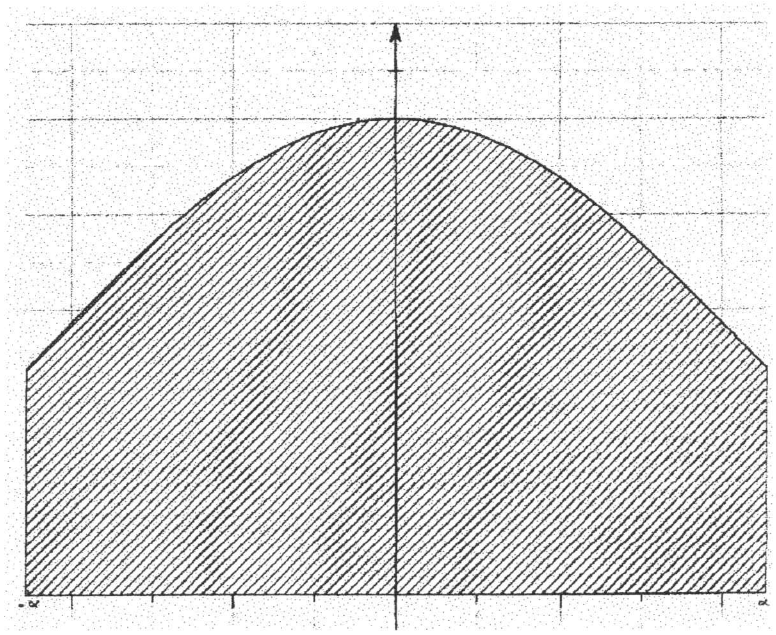


[5] Les illustrations proposées ne sont pas de Lagrange qui décrit seulement ce qui se passerait.

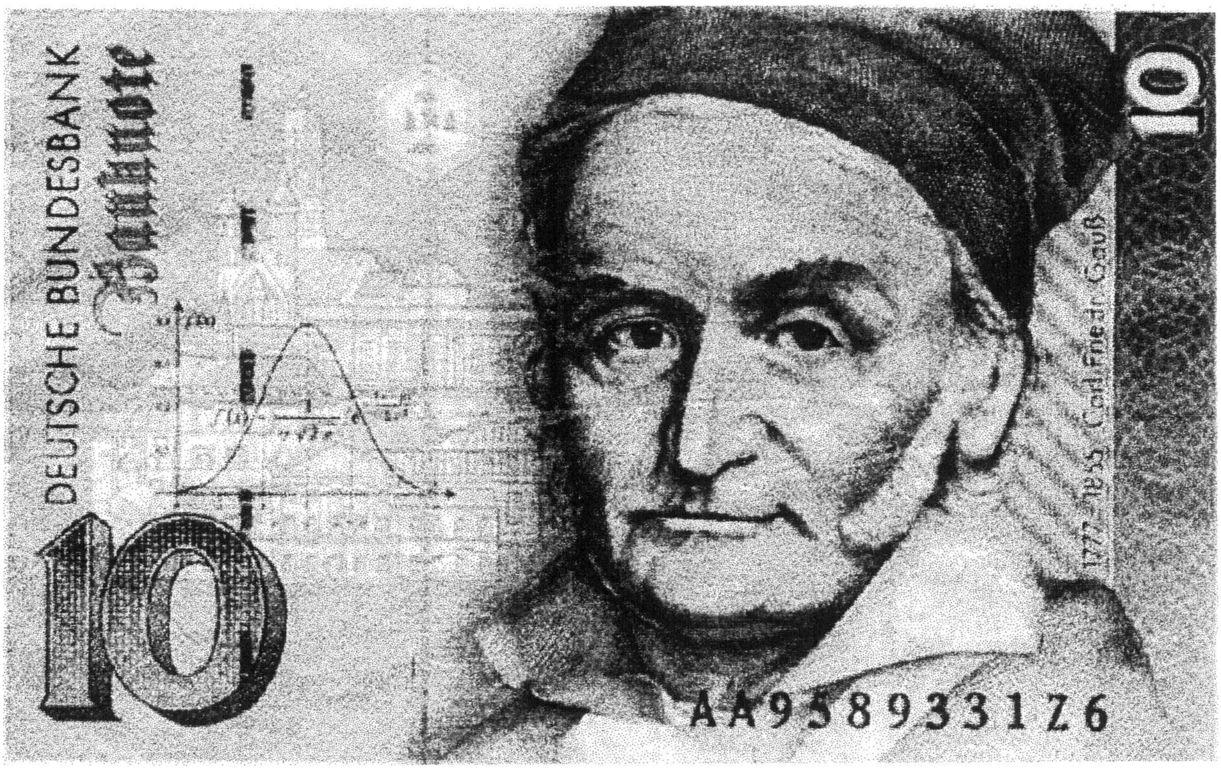
Alors, prolongeant au continu les résultats précédents, Lagrange pourra affirmer que la correction à faire au résultat moyen sera l'abscisse du centre de gravité de l'aire totale hachurée ci-dessus.

Un cas particulier : si la courbe est symétrique par rapport à l'axe des ordonnées, alors la correction est nulle.

"Ce cas a lieu toutes les fois que les erreurs peuvent être également positives et négatives."



[6] Le problème de la courbe de facilité des erreurs, ou loi des erreurs, sera examiné dans les chapitres suivants. Il s'agit effectivement de développer un modèle.



Carl Friedrich Gauss

3 - De la détermination de la "vraie valeur"

1 Les "vraies de vrai" mesures :

Vous êtes vous demandés un jour, combien mesurait « exactement » votre salle de cours ? Les « vraies de vrai mesures » ?

Vous penserez sûrement que cela a peu d'importance, et vous aurez raison. Mais si vous étiez un astronome et que vous vouliez mesurer la hauteur d'un astre[1] par exemple ... le problème, un peu plus délicat reconnaissons le, serait cependant du même ordre : connaître la « vraie valeur », alors qu'on ne dispose que d'instruments de mesure qui, même s'ils sont le plus précis possible, sont imparfaits. Ces problèmes se sont réellement posés et ils étaient d'une extrême importance.

Plaçons nous au milieu du XVIII° siècle. Depuis Copernic, Kepler et surtout Newton, l'astronomie est devenue une science étayée par une théorie mathématique. Mais le travail de l'astronome s'appuie sur des mesures, et si celles-ci sont entachées d'erreurs, même si la théorie est bonne, le résultat ne sera pas bon, ou au moins il sera contesté. Par exemple Clairaut et Lalande se trompent d'un mois sur la prévision du passage à la périhélie de la comète de Halley en 1759.

Une des grandes questions de cette époque est celle de la forme de la terre. Newton a « calculé » qu'elle devait être aplatie aux pôles. Des scientifiques français ne sont pas convaincus. Leurs mesures pour l'élaboration de la toute nouvelle carte de France sembleraient prouver que la terre est aplatie à l'équateur. Il sera donc décidé d'organiser deux expéditions, l'une à l'équateur, l'autre en Laponie, pour que les mesures sur le terrain donnent la « vraie » réponse. Au delà de l'aventure scientifique extraordinaire[2], ce sera une occasion de prendre conscience de la difficulté de mesurer, de se rendre à l'évidence qu'il est

---

[1] Vous pourrez lire utilement par exemple à ce sujet la brochure : Mesurer la terre aussi bien que le ciel.

[2] Nous conseillons à ce sujet la lecture du merveilleux livre de Florence Trystram : Le procès des étoiles.

impossible, expérimentalement, d'obtenir un résultat dont on puisse affirmer que c'est la « vérité », et de se poser la question : puisqu'on ne peut avoir la « vraie » valeur, comment l'approcher le mieux possible, et comment avoir une idée de l'erreur commise.

S'intéressant ainsi aux erreurs, les scientifiques essaient de trouver les moyens de combiner de nombreuses mesures d'un même phénomène, qui réduisent le plus possible l'erreur finale.

Pour en avoir le cœur net, revenons à notre salle de classe. A votre avis, si deux élèves mesurent la longueur de la salle, par exemple avec le même « mètre », trouveront-ils la même longueur ? Et s'ils disposent de deux instruments différents ? A plusieurs, essayez de réunir le plus d'instruments possibles : mètres rubans, mètres en bois articulés, chaîne d'arpenteur ..., et mesurez. Comparez les mesures obtenues (il en faudrait au minimum une cinquantaine), et essayez de répondre à la question : quelle est la longueur exacte de notre salle de classe ? Pour cela, vous pouvez répertorier les instruments dans lesquels vous avez le plus de confiance, et éventuellement, éliminer des mesures si elles vous semblent totalement farfelues.

Pensez-vous maintenant pouvoir donner une mesure, sans doute très proche de la longueur exacte ?

Nous la nommerons L.

Disposez alors toutes vos mesures dans un tableau et calculez les « erreurs » (en plus ou en moins) pour chacune d'elles, par rapport à cette valeur L.

Mesures effectuées L_i	L_1	L_2 ....
Différences $L_i - L$ (= erreurs évaluées)		

Dressons maintenant une sorte de tableau statistique des erreurs.

Erreurs (par ordre croissant)	
effectifs	

Par exemple, si vous avez trouvé 4 fois une erreur de – 0,01 m, vous placerez sur la première ligne – 0,01, et sur la deuxième ligne 4.

Reportez alors les éléments de ce dernier tableau sur un graphique, en prenant en abscisse les erreurs, et en ordonnée les effectifs.

En reliant les points obtenus, vous obtiendrez une courbe appelée la courbe des erreurs.

C'est ce que les mathématiciens du XVIIIe° siècle ont commencé à étudier : les courbes d'erreurs.

Pierre-Simon De Laplace

Bien sûr le plus difficile reste à faire, estimer la "vraie" valeur L, puisque c'est l'inconnue et que tout dépend d'elle. C'est donc la première question que se sont posés les scientifiques.

Voici par exemple l'article "milieu" du supplément à l'Encyclopédie de Diderot et D'Alembert, rédigé par Jean III Bernoulli, vers 1770 :

> " Quand on a fait plusieurs observations d'un même phénomène, et que les résultats ne sont pas tout à fait d'accord entr'eux, on est sûr que ces observations sont toutes, ou au moins en partie peu exactes, de quelque source que l'erreur puisse provenir ; on a coutume de prendre le milieu entre tous les résultats, parce que de cette manière les différentes erreurs se répartissant également dans toutes les observations, l'erreur qui peut se trouver dans le résultat moyen devient aussi moyenne entre toutes les erreurs. Il n'est pas douteux que cette pratique ne soit très utile pour déterminer l'incertitude qui naît de l'imperfection des instruments et des erreurs inévitables des observations ; mais il est aisé de s'apercevoir qu'elle ne la diminue pas autant qu'on le désirerait, et qu'elle est susceptible à plus d'un égard d'être perfectionnée, parce qu'en prenant simplement le milieu arithmétique, on ne tient pas compte du plus ou moins de probabilité de l'exactitude des observations, des différents degrés d'habileté des observateurs, etc. Différents grands géomètres ont entrepris cette recherche utile, ils l'ont considérée sous différents points de vue, et l'ont traitée plus ou moins en détail ; il est fort à souhaiter que les astronomes, les physiciens et généralement tous les observateurs profitent des résultats de ces recherches dans la discussion de leurs observations."

Vous remarquerez que J. Bernoulli ne veut pas a priori donner de forme mathématique à ce qu'il nomme "milieu". Il voudrait cependant une valeur optimale et signale qu'il faudrait pouvoir tenir compte du "plus ou moins de probabilité de l'exactitude des observations".

Le "milieu " L que vous avez choisi plus haut est peut-être la moyenne de vos mesures ; nous sommes en effet maintenant très habitués à utiliser les moyennes. En fait, cela ne va pas de soi. Lors de l'expédition à l'Équateur, par exemple, les mesures sur le terrain, faites par La Condamine et Bouguer étaient peu concordantes. La Condamine proposa alors à Bouguer de faire la moyenne des mesures (qui avaient été faites avec le plus grand soin). Bouguer s'y refusa longtemps, estimant que ce serait une "tricherie".

Supposant qu'un milieu ait été choisi, la deuxième étape est l'estimation de l'erreur, par rapport à la "vraie" valeur toujours inconnue. Ceci sera fait en utilisant ce qui sera appelé la distribution des erreurs.

Très rapidement on va chercher un modèle pour cette distribution, qui devra tenir compte des contraintes "naturelles "suivantes, signalées par Galilée :

Les observations sont distribuées de façon symétrique autour de la "vraie" valeur ; c'est-à-dire que les erreurs sont distribuées de façon symétrique autour de zéro. (A priori on fait autant d'erreurs dans un sens que dans l'autre).

Les petites erreurs arrivent plus fréquemment que les grandes erreurs. (Parce que l'observateur est consciencieux, donc il ne devrait pas faire beaucoup de grosses erreurs).

Gauss exprimera ces contraintes dans sa "Méthode des moindres carrés" en 1821 :

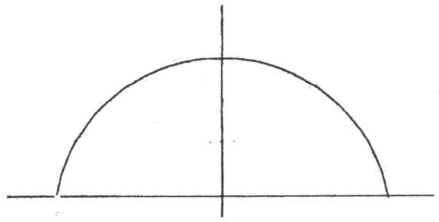
> "Les erreurs qui, dans des observations de même espèce, proviennent d'une cause simple et déterminée se trouvent renfermées entre certaines limites (...)
>
> Désignons par F(x) la facilité relative d'une erreur x : on doit entendre par là, à cause de la continuité des erreurs, que F(x)dx est la probabilité que l'erreur soit comprise entre les limites x et x + dx. Il n'est pas possible, en général, d'assigner la forme de la fonction F, et l'on peut même affirmer que cette fonction ne sera jamais connue dans la pratique. On peut néanmoins établir plusieurs caractères généraux qu'elle doit nécessairement présenter : (...) Elle s'annule pour toutes les valeurs de x non comprises entre les valeurs extrêmes. Pour toute valeur comprise entre ces limites, la fonction est positive. (...) Dans la plupart des cas, les erreurs égales et de signes contraires seront également probables.(...) Enfin, comme les petites erreurs sont plus facilement commises que les grandes, F(x) sera en général maximum pour x = 0 et diminuera sans cesse lorsque x croîtra."

Votre courbe d'erreur obéit-elle à ces contraintes ?

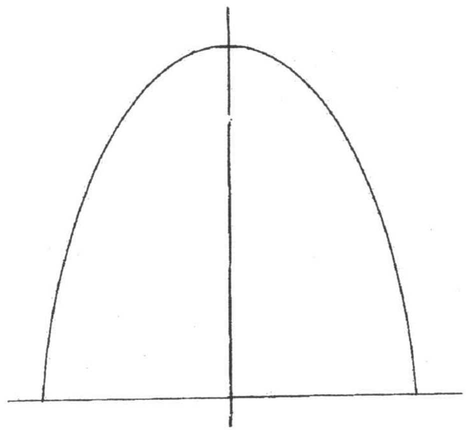
Dans le catalogue des courbes fonctionnelles que vous connaissez, y en a -t-il qui obéissent à ces contraintes ?

Plusieurs modèles ont été proposés. Citons quelques exemples.

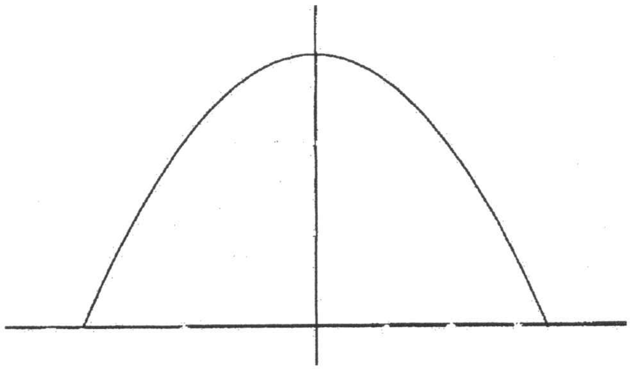
Daniel Bernoulli en a proposé trois : en arc de cercle, d'ellipse ou de parabole.



Arc de cercle



Arc d'ellipse



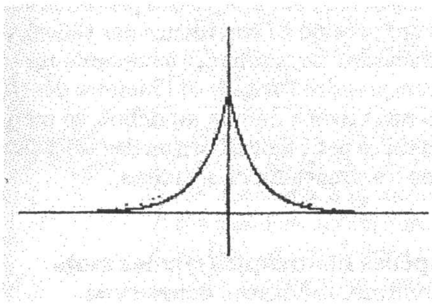
Arc de parabole

Daniel Bernoulli

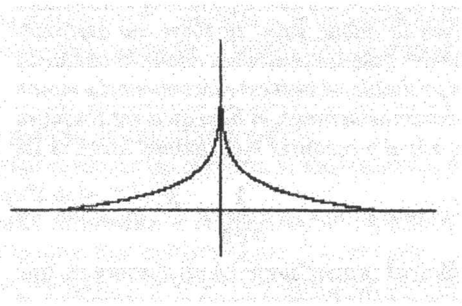
Lagrange envisagea une distribution cosinusoïdale.

Les plus avancés parmi vous pourront comprendre les modèles proposés par Laplace, qui font appel aux fonctions logarithmes et exponentielles.

Laplace propose en effet deux modèles (en 1774 et en 1777). Un modèle en $e^{-k|x|}$, sur $]-\infty ; +\infty[$, puis un modèle en $\log\frac{a}{|x|}$ sur $]-a ; a[$.



distribution en $e^{-k|x|}$,



distribution en $\log\frac{a}{|x|}$

Toutes les courbes proposées ci-dessus ont les caractéristiques demandées par Gauss.

Il va s'agir en fait de choisir un modèle de distribution des erreurs qui sera lié au choix du "milieu" ; le meilleur milieu sera celui qui minimisera les erreurs.

La première idée est de minimiser la somme des écarts à la vraie valeur.

Revenons à notre salle de classe. Vous ne connaissez toujours pas la vraie longueur. Nous la nommerons *l*.

Approximez *l* par la médiane de vos longueurs L_1, L_2, L_3 …. La médiane est un "milieu" possible.

Calculez d'une part la somme $|L_1 - l| + |L_2 - l| + |L_3 - l| + \ldots$, puis d'autre part la somme : $(L_1 - l)^2 + (L_2 - l)^2 + (L_3 - l)^2 + \ldots$

Approximez maintenant *l* par la moyenne arithmétique des longueurs L_1, L_2, L_3 ….Et calculez les mêmes sommes que précédemment.

Pour chacune de ces sommes quel est le meilleur milieu ? C'est-à-dire, quelle est la somme la plus petite ?

On peut établir que le choix de la médiane comme milieu minimise la somme des écarts, alors que celui de la moyenne minimise la somme des carrés.

Gauss choisira de minimiser la somme des carrés des écarts à la vraie valeur, donc de privilégier la moyenne.

Il peut sembler étrange a priori de minimiser la somme des carrés, mais, d'une part le calcul en était beaucoup plus simple, d'autre part, la moyenne était utilisée de manière empirique depuis des années.

Le terme "médiane" ne sera introduit qu'en 1843, par Cournot, dans "Exposition de la théorie des chances et des probabilités". En 1757, R. J. Boscovitch avait été un des premiers à souligner l'intérêt de cette valeur centrale.

Ainsi, le choix de la minimisation de la somme des carrés des écarts va s'imposer. Citons Legendre (Nouvelles méthodes pour la détermination des orbites des comètes - 1805).

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

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

Ce modèle, appelé *loi de Gauss* (ou *loi de Laplace-Gauss*) va s'imposer comme étant **La** loi de distribution des erreurs.

Gauss prend comme hypothèse que : "*Si une quantité a été déterminée par plusieurs observations directes, faites dans les mêmes circonstances et avec le même soin, la moyenne* *arithmétique des valeurs observées donne la valeur la plus probable, si ce n'est exactement du moins approximativement, si bien qu'il est toujours plus sûr d'y recourir"*

Il obtient ainsi la loi des erreurs : $\Phi(x) = \frac{1}{h\sqrt{\pi}} e^{\frac{-x^2}{h^2}}$, où h est un coefficient proportionnel à un nombre m, qu'il appelle "l'erreur moyenne à craindre", qui sera appelé en 1893 "standard deviation", par Karl Pearson. Il le note σ et c'est notre écart-type.

Karl Pearson

Vous aurez compris, nous l'espérons, au terme de cette histoire, que les choix opérés ont été quelque peu arbitraires. On ne connaît toujours pas la "vraie valeur". On minimise le mieux possible les erreurs, mais on aurait pu les minimiser autrement.

Dans le texte suivant[3], Gauss lui-même met très bien en évidence l'arbitraire de ses choix.

"La question qui nous occupe a, dans sa nature même, quelque chose de vague et ne peut être bien précisée que par un principe jusqu'à un certain point arbitraire. La détermination d'une grandeur par l'observation peut se comparer, avec quelque justesse, à un jeu dans lequel il y aurait une perte à craindre et aucun gain à espérer : chaque erreur commise étant assimilée à une perte que l'on fait, la crainte relative à un pareil jeu doit s'exprimer par la perte probable, c'est-à-dire par la somme des produits des diverses pertes possibles par leurs probabilités respectives. Mais quelle perte doit-on assimiler à une erreur déterminée ? C'est ce qui n'est pas clair en soi ; cette détermination dépend en partie de notre volonté. Il est évident, d'abord, que la perte ne doit pas être regardée comme proportionnelle à l'erreur commise ; car, dans cette hypothèse, une erreur positive représentant une perte, l'erreur négative devrait être regardée comme un gain : la grandeur de la perte doit, au contraire, s'évaluer par une fonction de l'erreur dont la valeur soit toujours positive. Parmi le nombre infini de fonctions qui remplissent cette condition, il semble naturel de choisir la plus simple, qui est, sans contredit, le carré de l'erreur, et, de cette manière, nous sommes conduits au principe proposé plus haut.

Laplace a considéré la question d'une manière analogue, mais en adoptant, pour mesure de la perte, l'erreur elle-même prise positivement. Cette hypothèse, si nous ne nous faisons pas d'illusion, n'est pas moins arbitraire que la nôtre : faut-il, en effet, regarder une erreur double comme plus ou moins regrettable qu'une erreur simple répétée deux fois, et faut-il, par suite, lui assigner une importance double ou plus que double ? C'est une question qui n'est pas claire, et sur laquelle les arguments mathématiques n'ont aucune prise ; chacun doit la résoudre à son gré. On ne peut nier pourtant que l'hypothèse de Laplace ne s'écarte de la loi de continuité et ne soit, par conséquent, moins propre à une étude analytique ; la nôtre, au contraire, se recommande par la généralité et la simplicité de ses conséquences.

Vous êtes maintenant dans un monde de statistiques, régi par des lois arbitraires, faute de mieux. Et cela ne fonctionne pas trop mal. Il faut cependant toujours conserver son sens

[3] Méthode des moindres carrés, 1821.

critique et garder par exemple à l'esprit cette boutade rapportée par Poincaré : "*Tout le monde y croit cependant, me disait un jour M. Lippman, car les expérimentateurs s'imaginent que c'est un théorème de mathématique, et les mathématiciens que c'est un fait expérimental.*"
La moyenne cependant a pris un rôle prépondérant dans la recherche des "vraies valeurs" des phénomènes, à tel point que certains ont eu l'idée d'appliquer ces résultats à d'autres domaines.

II Les moyennes "subjectives" et les statistiques sociales :

Si vous pensez à vos premiers contacts avec la notion de moyenne, vous songerez fort probablement, plutôt à la moyenne des notes, ou à la moyenne des tailles des élèves de la classe, que le professeur de mathématiques vous aura souvent proposé de calculer.
S'agit-il du même problème que précédemment ? Y aurait-il une "vraie" note, une "vraie" taille ? la réponse sera non, mais y réfléchissant est-ce si évident ? N'y aurait-il pas une notion de normalité qui se dégagerait ?
Il n'est pas très glorieux finalement de n'être que moyen. En particulier dans le cas des notes, qui évaluent certaines aptitudes intellectuelles, on calculera la moyenne, pour constater, si possible, qu'on est au-dessus.
Cependant, les méthodes pour le calcul de ces moyennes sont exactement les mêmes que pour déterminer les "vraies valeurs". Il y a donc lieu d'examiner si ces moyennes d'un autre type, qui donnent des renseignements d'un autre ordre, ont quelques relations avec les précédentes.
L'un de ceux qui va contribuer à introduire les méthodes statistiques dans les questions de sociologie est le mathématicien et astronome belge, Quételet.

"L'homme que je considère ici est, dans la société, l'analogue du centre de gravité dans les corps ; il est la moyenne autour de laquelle oscillent les éléments sociaux : ce sera si l'on veut un être fictif pour qui toutes les choses se passeront conformément aux résultats moyens obtenus pour la société. Si l'on cherche à établir, en quelque sorte, les bases d'une physique sociale, c'est lui qu'on doit considérer, sans s'arrêter aux cas particuliers ni aux anomalies."

Pour expliquer l'analogie entre la moyenne d'une série de n mesures d'un même objet (ce qu'on appelle la moyenne objective) et la moyenne d'une série de mesures de n objets différents (ce qu'on appelle moyenne subjective), il tient le raisonnement suivant[4] :

Si on fait mille mesures du tour de poitrine d'une statue, et que l'on représente graphiquement la série de mesures, on obtient une courbe de Gauss ;

Si, au lieu de mesurer mille fois la statue, on demande à mille sculpteurs de la copier, les mesures obtenues sur le tour de poitrine des mille statues suivront aussi une loi de Gauss ;

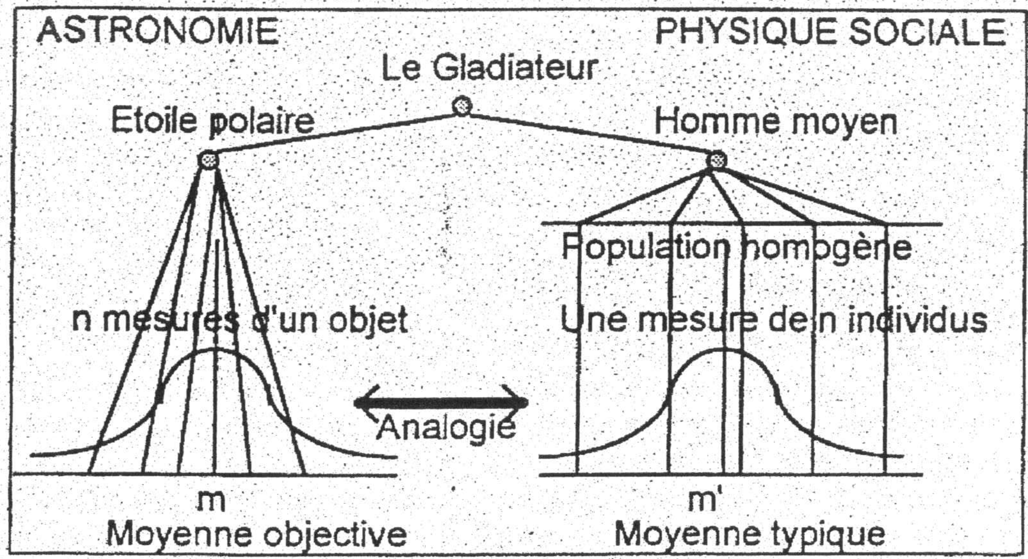
Maintenant, si, au lieu de mesurer mille statues, on mesure le tour de poitrine de mille individus, où est la différence ? on obtiendra aussi, nécessairement, une courbe de Gauss.

Voici le passage clé de sa lettre :

"Je demande maintenant s'il serait exagéré de parier un contre un qu'une personne peu exercée à prendre des mesures sur le corps humain va se tromper de 33 millimètres environ, en mesurant une poitrine de plus d'un mètre de circonférence ? Eh bien, en admettant cette erreur probable, 5738 mesures prises sur une même personne ne se regrouperaient certainement pas avec plus de régularité, quant à l'ordre de grandeur, que les 5738 mesures prises sur les soldats écossais. Et si l'on connaît les deux séries de mesures sans les avoir désignées d'une manière particulière, nous serions très embarrassés de dire quelle série a été prise sur 5738 soldats différents, et quelle série a été obtenue sur une seule et même personne, avec moins d'habitude et des moyens plus grossiers.

L'exemple que je viens de citer mérite, je crois, toute notre attention : il nous montre que les choses se passent absolument comme s les poitrines qui ont été mesurées avaient été modelées sur un même type, sur un même individu, idéal si l'on veut, mais dont nous pouvons saisir les proportions par une expérience suffisamment prolongée. Si telle n'était pas la loi de la nature, les mesures ne se regrouperaient pas malgré leurs défectuosités, avec l'étonnante symétrie que leur assigne la loi de possibilité."

[4] Quételet a entretenu une correspondance avec le Duc de Saxe-Cobourg, sorte de "cours de probabilité par correspondance". C'est dans la lettre XX qu'apparaît ce raisonnement.



Le transport analogique de la moyenne

Quételet avait observé que les distributions "en chapeau de gendarme" apparaissaient non seulement en astronomie, mais aussi dans d'autres mesures, comme celle des tailles des conscrits. Au bout de quelque temps, il vint à imaginer un "homme moyen", qui serait le plus équilibré possible, mais sans qualité distinctive, donc représentatif de n'importe qui[5].

Exercice :

Nous vous proposons d'établir pour votre classe, ou un petit groupe d'élève, d'évaluer la taille moyenne, la pointure moyenne, la longueur moyenne des bras, et d'imaginer quel est l'allure de l'élève moyen de la classe. Cela a-t-il beaucoup de sens ?

Ce sont les Bertillon[6] (père et fils), qui, à la fin du XIX° siècle, distingueront la "moyenne" objective, celle de la théorie des erreurs, correspondant à un objet réel, de la moyenne

---

[5] Nous vous conseillons, pour avoir une illustration philosophique et littéraire de "l'homme moyen", la lecture du roman de Robert Musil "L'homme sans qualité".

[6] Adolphe (1821-1883) et Alphonse (1853-1914) Bertillon sont considérés comme les inventeurs de l'anthropométrie judiciaire.

subjective, celle de l'homme moyen, qui fonde un objet fictif nouveau, sous la condition de "normalité". Du coup la loi de Gauss deviendra le plus souvent la loi "normale".

La statistique moderne, celle que l'on utilise par exemple dans les sciences sociales, naîtra en Angleterre. Il s'agit, dépassant Quételet, de le transformer complètement, en particulier sur la question de l'homme moyen. Ce sont les travaux de Galton, cousin de Darwin, qui sont à l'origine de cette statistique. La loi normale va maintenant être plutôt utilisée comme un outil pour étalonner les aptitudes. Et bien sûr, il s'agit d'améliorer l'espèce en favorisant ceux qui sont au-dessus de la moyenne. Il est facile de voir les glissements possibles vers des théories eugénistes.

Pour conclure sur cette histoire, nous pourrions dire que les statistiques actuelles sont évidemment héritières de ces divers courants. Mais, du moins si on se laisse porter par les apparences, la statistique est plutôt devenue une méthodologie, un ensemble de techniques, que le spécialiste sur son terrain appliquera, en interprétant les résultats dans le champ qui lui est propre. Il semble pourtant essentiel d'examiner sans cesse à quoi servent les outils statistiques, pourquoi et dans quel contexte philosophique ou social ils ont été pensés.

Les calculs mathématiques, en tant que tels et pris isolément, ne sont pas contestables. L'usage qu'on en fait demande un regard critique. Il faut reconnaître que cela ne semble pas si mal marcher. Une fois de plus les mathématiques serviraient à décrire et comprendre le réel. Il faut prendre garde cependant que ce regard sur le réel ne soit pas biaisé. Un regard sur la construction des concepts peut contribuer à nous préserver de cet écueil.

Francis Galton

4. Médiane ou moyenne, laquelle choisir ?

Comme nous l'avons vu, la médiane et la moyenne sont deux paramètres différents qui traduisent cependant tous deux l'idée de « milieu » d'une série statistique. Il n'est pas toujours facile de choisir lequel de ces deux paramètres va être le plus judicieux pour traduire cette idée de « milieu ».

Au 17ème siècle les frères Louis et Christian Huygens ont eu une discussion fort intéressante à ce sujet.

Christian Huygens

A cette époque (1662) l'anglais John Graunt, ayant fait des relevés statistiques sur la mortalité dans la ville de Londres, fait paraître une « table de mortalité ». Cette table est reproduite ci-dessous :

Nous avons trouvé que, sur 100 individus conçus et animés, 36 environ meurent avant l'âge de 6 ans et peut-être un seul est survivant à 76 ans. Comme il y a 7 décennies entre 6 et 76, nous avons recherché six nombres moyens proportionnels entre 64, nombre de survivants à 6 ans, et 1, celui qui survit à 76 ans, et nous trouvons que les nombres suivants sont pratiquement assez proches de la vérité, car les hommes ne meurent pas selon des proportions exactes ni en fractions[1].

Sur 100 individus, il meurt pendant les six premières années	36
les dix années suivantes ou 1ère décennie	24
la 2ème décennie	15
la 3ème décennie	9
la 4ème décennie	6
la suivante	4
la suivante	3
la suivante	2
la suivante	1

Il s'ensuit que, sur ces 100 individus conçus, il en survit

Au bout de	6 ans	64
	16 ans	40
	26 ans	25
	36 ans	16
	46 ans	10
	56 ans	6
	66 ans	3
	76 ans	1
	86 ans	0

La « petite table anglaise » de John GRAUNT extraite de
Natural & Political Observations Upon the Bills of Mortality… of the City of London (1662).
Traduction française par E. Vilquin. Ed. INED 1977.

[1] En faisant le calcul indiqué par J.Graunt on ne trouve pas exactement les nombres qu'il donne : la raison de la suite géométrique de 1er terme 64 et de 8ème terme 1 est de $1/\sqrt[7]{64}$, ce qui donne des calculs assez compliqués ; il semble qu'il ait pris 5/8 comme raison.

Louis Huygens a connaissance de cette table et l'exploite de la manière suivante : il calcule *la durée de vie moyenne* d'un homme et trouve 18 ans et 2 mois, c'est à dire que l'espérance de vie *à la naissance* est de 18 ans et 2 mois. Puis il calcule l'espérance de vie à 6 ans, à 16 ans, etc..., c'est à dire la durée de vie d'un homme qui est parvenu à l'âge de 6 ans, 16 ans, etc... Voici les résultats qu'il obtient :

Age	**Espérance de vie**
16 ans	20 ans 3 mois
26 ans	19 ans 4 mois
36 ans	17 ans 6 mois
46 ans	15 ans
56 ans	11 ans 8 mois
66 ans	8 ans 4 mois
76 ans	5 ans
86 ans	0

Il écrit alors à son frère Christian pour lui faire part de ses recherches.

> J'ay fait une table ces jours passez du temps qu'il reste à vivre à des personnes de toute sorte d'aage, de laquelle je vous envoie ici une copie, a fin que vous preniez la peine de faire un peu les mesmes supputations, et que nous puissions voir comment nos calculs s'accorderont. J'advoüe que j'ay eu assez de peine d'en venir à bout, mais à vous il n'en sera pas de mesmes, et les conséquences qui en resultent sont fort plaisantes et peuvent mesme estre utiles pour les constitutions des rentes à vie. La question est jusqu'à quel aage doibt vivre naturellement un enfant aussi tost qu'il est conceu. Puis un enfant de 6 ans, puis un de 16 ans, de 26, etc. Si vous y trouvez de la difficulté ou trop d'embaras, je m'offre à vous faire part de ma méthode, qui est asseuree, par la première occasion.
> Adieu.
> Selon mon calcul, vous vivrez environ jusqu'à l'aage 56 ans et demy. Et moy jusqu'à 55.
>
> *Extrait d'une lettre de Louis HUYGENS à son frère Christian (22 août 1669)*

La réponse de Christian ne se fait pas attendre. Loin de complimenter son frère sur ses calculs il met en doute leur bien-fondé et décide de faire ses propres calculs.

C'est beaucoup fait à vous d'avoir peu faire le calcul des aages, dont vous dites estre venu à bout. Mais a fin que ce calcul fust exact, il faudrait avoir une table qui marquast en année combien il meurt des personnes de 100 qu'on suppose, et il faut que vous l'ayez suppléée par quelque moyen comme j'en scay pour cela, ou autrement vous ne scauriez determiner au vray combien doibt vivre une personne de 6, 16 ou 26 ans, etc, et encore moins de quelque aage moyen entre ceuxlà, comme vous l'avez entrepris de vous et de moy. Je crois donc que vous n'en décidez qu'à peu près.

Ce que je puis conclure de certain par les données de la table, c'est que qui gagerait qu'un enfant nouveau-né (ou conceu comme vous dites, mais il me semble que l'Anglais ne parlait pas des conceus, car comment en peut-on tenir registre ?) vivra à 16 ans, prendroit le mauvais party et hazarderoit 4 contre 3. De même, qui gageroit qu'une personne de 16 ans vivra jusqu'à 36, il hazarde tout de mesme 4 contre 3[2].

J'ai envie de suppléer la table comme j'ay dit, et resoudre les problemes qu'on peut proposer en cette maniere qui est assez subtile. Vostre méthode ne sçauroit estre la mesme que la miene, et je seray bien aise de la voir. Adieu.

Extrait d'une lettre de Christian HUYGENS à son frère Louis (28 août 1669)

Ses résultats ne sont pas du tout les mêmes que ceux de son frère : pour lui l'espérance de vie à la naissance est de 11 ans ; quelqu'un ayant atteint l'âge de 6 ans peut espérer vivre jusqu'à 21 ans environ ; quelqu'un ayant atteint l'âge de 16 ans peut espérer vivre jusqu'à 31 ans ; etc...
Il écrit une nouvelle lettre à Louis dans laquelle il donne des arguments contre la méthode de celui-ci et où il lève le voile sur la sienne.

Je viens d'examiner vostre calcul des aages et de refaire le mien que j'avois perdu. Je voudrois que le vostre fust veritable, puis qu'il nous donne un peu plus de vie, mais il ne sert de rien de nous flatter ; *Scit nos Proserpina canos*[3], et elle ne s'arreste pas au compte que nous faisons. Vous concluez assez pres du vray, que les 100 personnes ont a faire ensemble 1822 ans de vie, mais il ne s'en suit pas que les 18 ans et 2 mois, qui viennent en divisant ce nombre par 100. soit l'aage de chasque personne creee ou conceue, ainsi que vous tenez pour certain. Prenons, par exemple, que les hommes soient encore plus foibles dans leur enfance qu'ils ne le sont, et que de 100 il en meure d'ordinaire 90 dans les premieres 6 annees, mais que ceux aussi qui surpassent cet aage soient des Nestors, et Mathusalems, et qu'ils vivent d'ordinaire jusqu'à 152 ans et 2 mois. Vous aurez pour les 100. le mesme nombre de 1822 ans, et cependant qui gageroit, qu'un enfant conceu parviendroit alors à l'aage de 6 ans seulement, auroit grand desavantage, puis que de 10 il n'y a qu'un qui y parvient.

Il ajoute un peu plus loin :

---

[2] En fait Christian se trompe : puisque sur 100 nouveau-nés 60 meurent avant 16 ans, c'est 60 contre 40, c'est à dire 3 contre 2 qu'il aurait fallu dire. De même pour la deuxième affirmation : sur 40 personnes parvenues à l'âge de 16 ans il en meurt 24 avant 36 ans : il aurait fallu dire 24 contre 16, c'est à dire encore 3 contre 2. D'ailleurs Christian rectifiera son erreur dans une lettre ultérieure (non citée ici).

[3] *Proserpine sait que nos cheveux sont blancs.* Proserpine (ou Perséphone chez les Grecs) fut enlevée par Pluton qui l'épousa et régna avec elle sur les Enfers.

> [...] vous demanderez comment je pourray determiner comme vous, combien il reste raisonnablement à vivre à une personne d'un aage proposé. Pour faire cela j'ay supplée la petite table angloise, sans pourtant m'embarrasser d'aucun calcul, mais en traçant une ligne courbe, sur la quelle avec le compas je mesure la vie de celuy qu'on veut, et je vois par exemple qu'à vostre aage de 38 ans, vous pouvez encore faire estat de 19 ans et 4 mois environ . Mais si vous vous amusez à faire appeler souvent des gens pour vous battre, il faut encore en retrancher quelque chose. Je vous envoieray la ligne de vie une autre fois avec la pratique d'icelle et mesme une table des vies à chaque aage d'annee en annee, qui ne me coustera guere.
>
> *Extrait d'une lettre de Christian HUYGENS à son frère Louis (21 novembre 1669)*

Comment se fait-il qu'à partir des mêmes données statistiques (la table de John Graunt) les frères Huygens n'arrivent pas aux mêmes résultats ? Ce sont tous deux des savants, qui ne font ni erreurs de calcul (en principe !) ni erreurs de raisonnement. Examinons donc leurs méthodes. (Les textes originaux sont donnés en annexe 1).

La table de John Graunt peut se résumer dans le tableau suivant :

Durée de vie (en années)	[0,6[	[6,16[	[16,26[	[26,36[	[36,46[	[46,56[	[56,66[	[66,76[	[76,86[
effectif	36	24	15	9	6	4	3	2	1

Pour déterminer *la durée de vie moyenne (ou l'espérance de vie à la naissance)* Louis Huygens fait le calcul suivant :

$$\frac{3\times36+11\times24+21\times15+31\times9+41\times6+51\times4+61\times3+71\times2+81\times1}{100}=18{,}22$$

c'est-à-dire 18 ans et environ 2 mois.

On reconnaît là le calcul de la **moyenne arithmétique pondérée** d'une série statistique dont les valeurs du caractère ont été regroupées en classes (on prend pour valeur du caractère le centre de la classe).

Pour déterminer la durée de vie de quelqu'un ayant déjà vécu jusqu'à l'âge de 6 ans, il refait le même calcul, mais en supprimant la première colonne du tableau :

$$\frac{11\times24+21\times15+31\times9+41\times6+51\times4+61\times3+71\times2+81\times1}{64}=26{,}78$$

soit 26 ans et environ 10 mois

Et ainsi de suite.

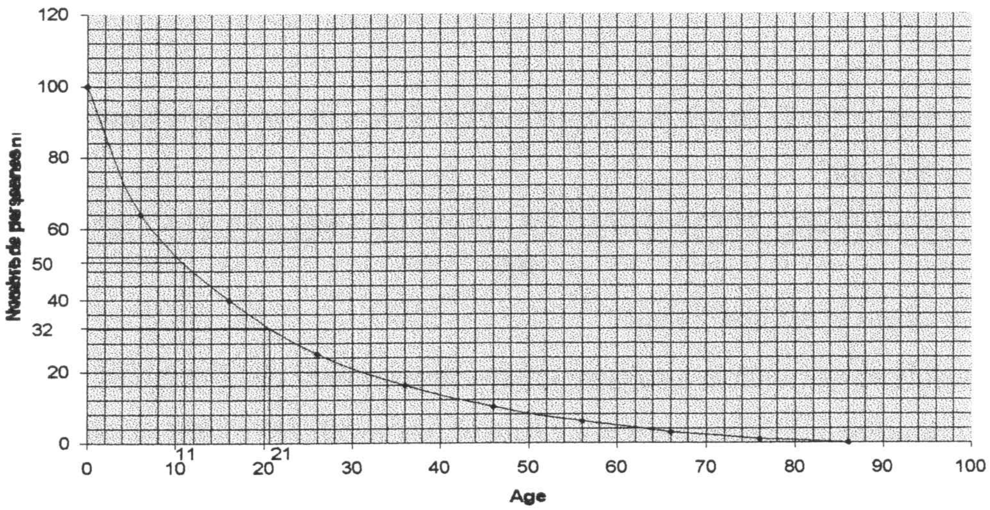
Louis Huygens calcule des **moyennes.**

Que fait Christian ? Il cherche à partir de quel âge il ne reste de vivant que la **moitié de l'effectif total** et trouve environ 11 ans : de nos jours, on dirait que la *durée de vie probable* est de 11 ans. De même, pour déterminer la durée de vie de quelqu'un ayant déjà vécu jusqu'à l'âge de 6 ans, il cherche à partir de quel âge ne survit que la moitié de l'effectif restant et il trouve environ 21 ans et ainsi de suite.

Christian Huygens calcule des **médianes**. La durée de vie probable à la naissance est la **médiane** de la série statistique.

Pour déterminer ces médianes Christian Huygens ne fait pas de calcul : il a représenté graphiquement la série statistique en mettant en abscisse l'âge et en ordonnée le nombre de personnes ayant atteint cet âge. Ainsi, pour déterminer la médiane il lit sur le graphique à quelle abscisse correspond un effectif de 50 ; puis pour déterminer l'âge auquel parviendra une personne ayant déjà atteint l'âge de 6 ans, il cherche l'abscisse du point d'ordonnée 32 (32 est la moitié de 64, nombre de personnes encore en vie au bout de 6 ans).

Voir le graphique ci-dessous.



Qui a raison ? Qui a tort ? Laissons à Christian Huygens le soin de conclure.

> Le calcul que je vous ay envoiè vous aura embarassé sans doute, au quel ayant songé depuis, et aussi au vostre, je trouve que nous avons tous deux raison en prenant la chose en different sens. Vous donnez à un enfant conceu 18 ans et 2 _ mois de vie, et il est vray que son esperance vaut autant que cela. Cependant il n'est pas apparent qu'il vivra si longtemps, car il est beaucoup plus apparent qu'il mourra devant ce terme.
> [...] Ce sont donc deux choses différentes que l'esperance ou la valeur de l'aage future d'une personne, et l'aage auquel il y a egale apparence qu'il parviendra ou ne parviendra pas. Le premier est pour regler les rentes à vie, et l'autre pour les gageures. Je verray si vous avez faict la mesme distinction. Cependant votre methode est fort belle et subtilement trouvée.
>
> *Extrait d'une lettre de Christian HUYGENS à son frère Louis (29 novembre 1669)*

Exercice : 1. A l'époque de cette correspondance entre les frères Huygens on sait que Louis a 38 ans (cf. lettre du 21/11/1669). D'après ses calculs, il dit qu'il vivra jusqu'à 55 ans et son frère Christian jusqu'à 56 ans et demi.

a. Retrouver le calcul que Louis a fait pour obtenir 55 ans.

b. Reprendre cette méthode pour connaître l'âge de Christian au moment de la correspondance.

2. En utilisant la méthode de Christian, déterminer l'âge probable du décès de chacun des deux frères. (Il semble bien que Christian se soit encore trompé ...).

ANNEXE 1

Extrait d'une lettre de Louis HUYGENS à son frère Christian (30 octobre 1669)

J'advoüe que mon calcul des aages n'est pas tout à fait juste, mais il y a si peu à dire que cela n'est aucunement considerable, et d'autant moins que la table Angloise, sur laquelle nous nous fondons, n'est pas dans cette derniere justesse aussi bien, mais comme dit cet Autheur, « those numbers are pratically neere enough to the truth, for men doe not die in exact proportions nor in fractions ». Voylà donc la methode dont je me suis servy. Je compte premierement les années que toutes ces 100. personnes ensemble doivent avoir vescu, qui font en tout 1822. années, ce que vous verrez prouvé dans la page qui suit.

Les 36 personnes qui meurent au dessoubs de 6. ans ont vescu l'un portant l'autre, 3. ans, qui fait	108. ans
Les 24. qui meurent entre 6. et 16. ont vescu l'un portant l'autre 11. ans, qui fait	264.
Les 15. qui meurent entre 16. et 26. ont vescu 21. ans, qui fait	315.
les 9. entre 26. et 36. ont vescu 31. ans, qui fait	279.
les 6. entre 36. et 46. ont vescu 41. ans, qui fait	246.
les 4. entre 46. et 56. ont vescu 51. ans, qui fait	204.
les 3. entre 56. et 66. ont vescu 61. ans, qui fait	183.
les 2. entre 66. et 76. ont vescu 71. ans, qui fait	142.
Et l'un qui meurt entre 76. et 86. a vescu 81. ans	81.
Somma	1822 ans

Ces 1 822. ans partagez esgalement entre 100. personnes il vient pour chacun 18. ans et environ 2. mois, qui est l'aage de chaque personne creee ou conceüe, l'une portant l'autre. Car notez en passant que c'est des personnes conceües que l'Anglois parle, et il en peut bien tenir registre aussi bien que de ceux qui sont nez, parce que les fausses couches entrent aussi dans ses observations. Or pour venir à notre compte et spécifier combien il reste de vie à chaque personne d'un tel ou tel aage, voylà comme je fay.

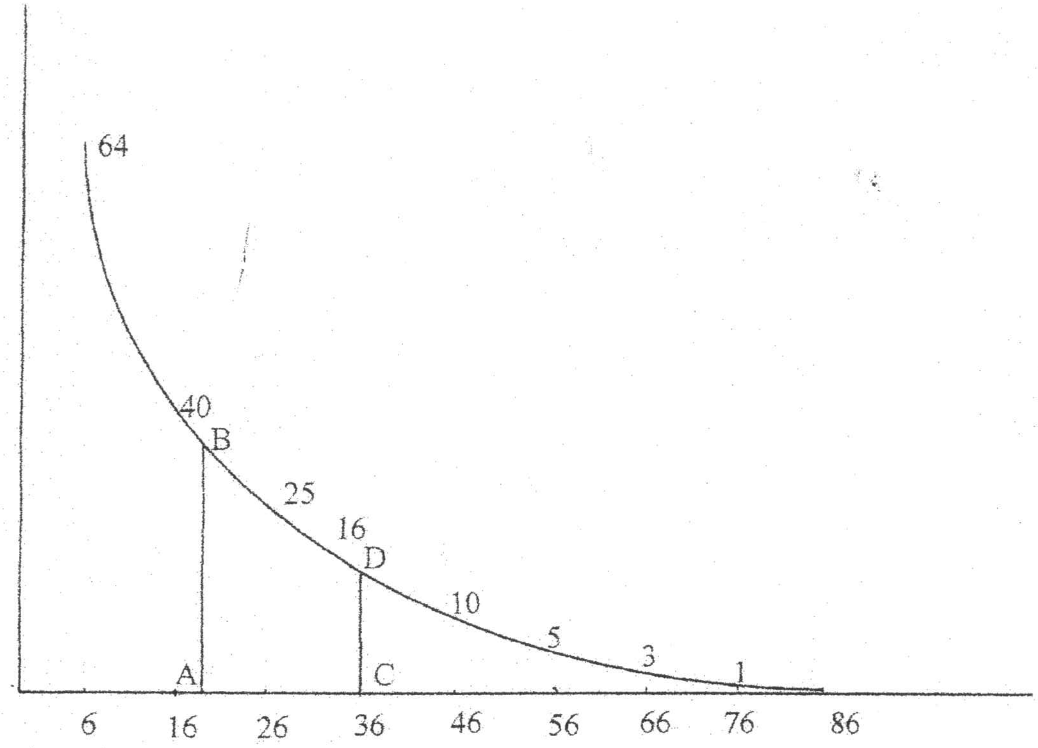
J'oste premièrement les 108. ans (qui est l'aage des 36. enfants qui meurent au dessoubs des 6. ans) de tout ce nombre de 1 822. ans ; reste 1 714. ans, lesquels doivent estre partagez entre les 64. personnes qui restent, ce qui fait pour chacun, c'est-à-dire pour chaque enfant de 6. ans, 26. ans et environ 10. mois, de sorte qu'il leur reste encor à vivre au susdit aage de 6. ans, 20. ans et 10. mois.

En suite ostez de ces 1714. ans, l'aage des 24. personnes qui meurent entre 6. et 16. (qui est 264. ans) il restera 1450. Lesquels se doivent partager entre les 40. personnes qui restent, ce qui fait pour chacun d'eux, c'est à dire pour chaque personne de 16. ans, 36. ans et 3. mois, de sorte qu'il

leur reste de vie	20. ans	3. mois.
Pour ceux de 26. il viendra 45. ans 4. mois, ou pour leur reste	19.	4.
Pour ceux de 36. 53. ans 6. mois ; pour leur reste	17.	6.
Pour ceux de 46. 61. ans. Pour leur reste	15.	0.
Pour ceux de 56. 67. ans 8. mois. Pour leur reste	11.	8.
Pour ceux de 66. 74. ans 4. mois. Pour leur reste	8.	4.
Pour ceux de 76. 81. ans. Pour leur reste	5.	0.
Pour ceux de 86. Rien	0.	

En suite de ce que dessus, je ne comprends pas la raison de vostre calcul de 4 contre 3 car, à mon advis, la partie est environ esgale lors qu'on gage qu'une personne de 6. ou une de 16. vivront encor 20 ans. J'attends donc vos raisons comme je vous ay envoyé les miennes.

Manuscrit conservé à la bibliothèque de Leyde (collection Huygens)



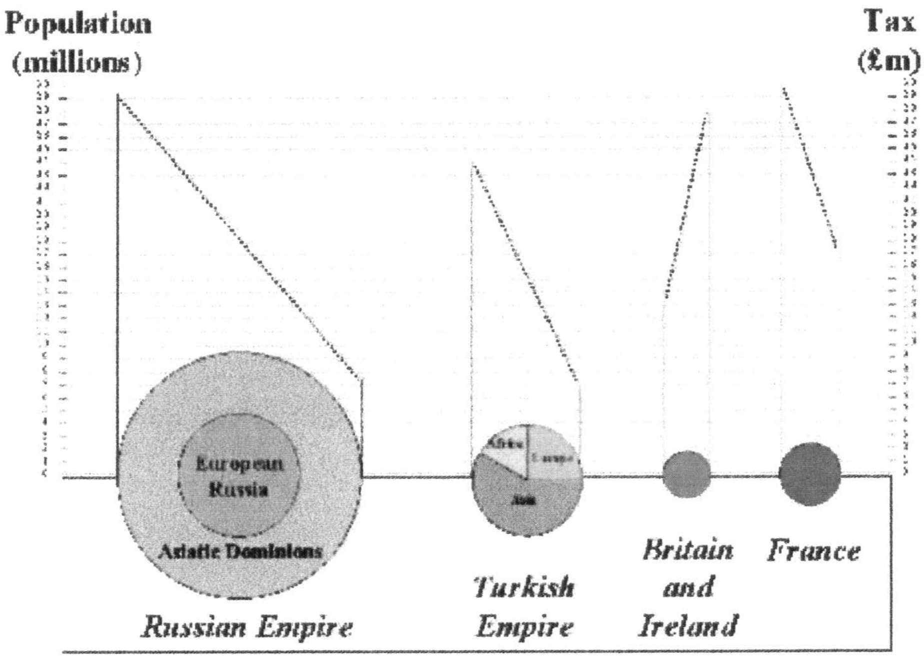
Sur la ligne droite d'embas sont marqués les aages des personnes et sur les 6 il y a une perpendiculaire de 64 parties parce que de 100 personnes selon la table angloise il en reste 64 à l'aage de 6 ans. Sur le 16 il y a une perpendiculaire de 40 parties parce qu'à l'aage de 16 ans il reste 40 personnes des 100 qui estoient conceues, et ainsi du reste. Et par tous les points ou bouts de ces perpendiculaires j'ay mené la ligne courbe 64, 40, 25 etc. Si je veux sçavoir maintenant combien il reste de personnes après les 20 années de 100 enfants conceus, je prens sur la ligne d'embas l'aage de 20 ans au point A d'ou ayant erigé une perpendiculaire qui rencontre la courbe en B, je dis que AB, qui pris sur l'eschelle d'embas fait presque 33 parties, est le nombre des personnes qui de 100 conceus atteignent l'aage de 20 ans, que si je veux sçavoir ensuite combien il reste raisonnablement à vivre à une personne de 20 ans par exemple, je prens la moitié de BA et l'ajuste en DC entre la courbe et la droite en sorte qu'elle soit perpendiculaire à la dernière. Et j'ay AC pour les annees qui restent à vivre à la dite personne, qui font pres de 16 ans, comme il paroit par les divisions dont chacune est une annee. La raison est, que la perpendiculaire DC estant la moitié de BA qui marquoit le nombre d'hommes qui restent des 100, 20 ans après la conception, à sçavoir 33, cette DC tombant sur 36 de la droite marquera qu'il reste la moitié de 33 c'est à dire 16 _ hommes après la 36^{e} année. Donc puis que des 33 personnes de 20 ans la moitié meurt d'ordinaire dans les prochains 16 ans, on peut gager avec egal avantage qu'une personne de 20 ans vivra encore 16 ans. On trouvera de mesme que la vie d'un enfant conceu doit estre taxee à 11 ans, au lieu que mon frere contait 18 et 2 mois.

Christian Huygens

Education reduces crime (1847)

	% difference in crime rate from average in counties of:	
	Worse education	Better education
More wealth	+ 9.2	- 29.4
Less wealth	+11.3	- 13.5

Source: F. G. P. Neison, BAAS, Oxford, 1847 (data for 1843/4)



Tableaux statistiques et graphiques de l'époque de Florence Nightingale.

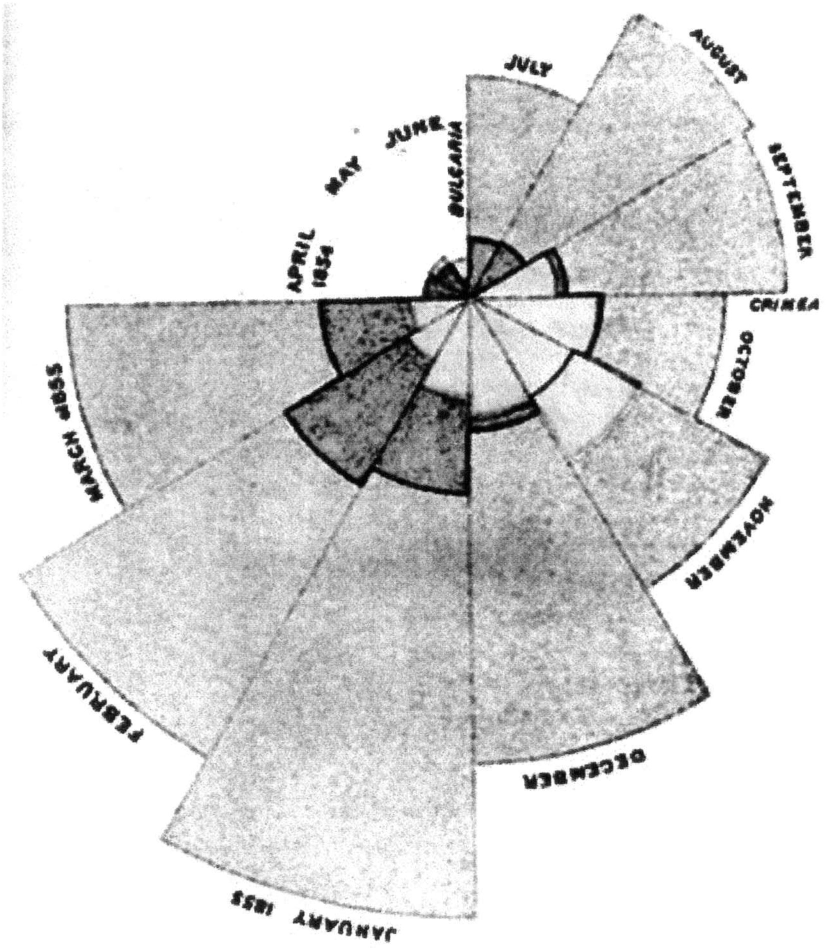
5. De l'utilité des graphiques en Statistique

I. Pour faire comprendre :

On a dit que Florence Nightingale avait été la première à utiliser des graphiques pour présenter des données statistiques. Ce n'est pas exact : le plus connu des pionniers en la matière est William Playfair qui publia en 1801 un graphique permettant de voir, par comparaison avec d'autres pays que les Anglais payaient plus d'impôts que les autres. Ce qui différencie Florence Nightingale de ses prédécesseurs est qu'elle utilise ses graphiques non seulement pour *montrer,* mais aussi pour *prouver* et ainsi faire comprendre la nécessité d'un changement .

Florence Nightingale

Le diagramme ci-dessous est un exemple de diagramme polaire inventé par Florence Nightingale - F. Nightingale l'appelle diagramme « en crête de coq » - pour donner une représentation graphique de la mortalité des soldats anglais durant la guerre de Crimée (1854-1856) et faire comprendre ainsi au Gouvernement la nécessité d'améliorer le niveau d'hygiène dans l'armée.



On doit lire ce diagramme de la façon suivante :

les aires des secteurs de couleur bleue, rouge et noire sont mesurées chacune à partir du centre du cercle ;

l'aire des secteurs de couleur bleue, mesurée à partir du centre du cercle, est proportionnelle au nombre de décès causés par des maladies contagieuses telles que le typhus, le choléra, contractées à l'hôpital ; celle des secteurs de couleur rouge, toujours mesurée à partir du centre du cercle, est proportionnelle au nombre de morts suite à des blessures reçues sur le champ de bataille ; et enfin celle des secteurs de couleur noire, mesurée aussi à partir du centre du cercle, est proportionnelle au nombre de décès dues à d'autres causes que celles citées précédemment.

Remarque importante : *les secteurs de couleur rouge et noire peuvent se recouvrir partiellement (Janvier 1855, Février 1855), exactement (Octobre 1854) ou même se recouvrir de telle sorte que le secteur rouge masque totalement le secteur noir (Novembre 1854 : la ligne noire marque la limite du secteur noir).*

Exercice : Sachant qu'en Janvier 1855 il y a eu 324 décès dus à des causes autres que les maladies contagieuses ou les blessures, calculer le nombre de décès dus aux maladies contagieuses en Janvier 1855, Février 1855 et Mars 1855. Quels pourcentages, par rapport au nombre total de décès, cela représente-t-il ?

II. Pour modéliser un phénomène :

Lorsqu'on dispose de données statistiques concernant un phénomène, on cherche à trouver un modèle mathématique qui lui corresponde. La représentation graphique de ces données peut être un moyen très simple d'y parvenir. Nous allons étudier quelques exemples.

Premier exemple : modèle affine

On fabrique en grande série une pièce dont une cote, exprimée en mm, doit se trouver dans l'intervalle de tolérance [51,8 ; 52,8]. En cours de fabrication, on prélève tous les quarts d'heure un échantillon pour lequel on calcule la valeur moyenne de cette cote. Le tableau suivant donne les résultats des deux premières heures de fonctionnement :

Durée x_i (en heures)	0,25	0,5	0,75	1	1,25	1,5	1,75	2
Moyenne y_i (en mm)	52,12	52,16	52,24	52,28	52,32	52,37	52,44	52,47

Les points de coordonnées (x_i, y_i) représentant ces données dans un repère orthogonal sont sensiblement alignés[1]. On peut donc considérer que le modèle associé est le modèle affine et il est alors très facile d'en donner les caractéristiques en cherchant une équation de la droite qui passe au plus près des points. *(On trouve, par exemple, $y = 0,2x + 52,07$)*

Lorsque les points représentant la série statistique sont alignés, c'est facile.

Deuxième exemple

Afin de déterminer le lien entre la distance d'arrêt d'une automobile et sa vitesse, on procède à l'expérience suivante : lorsque le pilote de l'automobile reçoit un signal sonore dans son casque, il doit arrêter son véhicule le plus rapidement possible ; au moment du top sonore on mesure la vitesse de l'automobile, puis on mesure la distance d'arrêt. Pour six expériences, à des vitesses différentes, on a obtenu les résultats suivants : (v_i est la vitesse au moment du top sonore, y_i est la distance d'arrêt).

v_i en $km.h^{-1}$	27	43	62	80	98	115
y_i en m	6,8	20,5	35,9	67,8	101,2	135,8

Ici les points de coordonnées (v_i, y_i) représentant ces données ne sont pas alignés. Il ne s'agit donc pas d'un modèle affine. Un œil un peu exercé doit pouvoir conjecturer que la courbe obtenue est une parabole mais ce n'est pas si simple de trouver son équation. N'y a-t-il pas moyen de « redresser » cette parabole ?

Faisons un nouveau graphique[2] : en ordonnée, on garde la même graduation ; en abscisse, on reporte le **carré** des réels . Si l'on représente la série sur ce nouveau graphique, on obtient des points (presque) alignés ! On peut alors, comme dans l'exemple 1, tracer la droite qui paraît la mieux à même de représenter la série et en trouver une équation.

[1] On trouvera le graphique en annexe 2.

On obtient, par exemple, $\frac{y}{X} = \frac{67,8}{80^2}$ (la droite passe par l'origine du repère) d'où y = 0,011X ; or $X = x^2$ et donc finalement la parabole obtenue a pour équation $y = 0,011\ x^2$

On a ainsi, grâce au graphique, trouvé un modèle mathématique, ce qui peut permettre de déterminer la distance d'arrêt d'un véhicule roulant à 50 $km.h^{-1}$ ou à 130 $km.^{-1}$ par exemple (*on trouve respectivement 27,5 m et 185,9 m).*

Troisième exemple : modèle exponentiel

On verse du thé très chaud (85° C) dans un bol à température ambiante. Le thé va se refroidir et on voudrait trouver un modèle mathématique pour ce phénomène. On mesure la température du thé toutes les minutes et on obtient les résultats suivants :

t (en minutes)	0	1	2	3	4	5	6	7	8
θ (en °C)	85	73	63	55	49	43	39	35	32

Si on reporte ces valeurs sur un graphique « classique », les points ne sont pas alignés. Mais si on utilise un papier sur lequel la graduation en ordonnée suit une échelle logarithmique, on obtient des points sensiblement alignés. Comme expliqué dans les exemples précédents, on peut tracer une droite passant au plus près des points et en donner une équation.

On trouve, par exemple, $\frac{Y - \ln 85}{x - 0} = \frac{\ln 39 - \ln 85}{6 - 0}$, d'où $Y = \frac{1}{6} x \ln\frac{39}{85} + \ln 85$ puis, en remplaçant Y par $\ln\theta$ et x par t, $\theta = 85\ e^{-0,13t}$.

Remarque : il existe dans le commerce du papier semi-logarithmique[3] *où l'une des graduations est faite à partir de la fonction logarithme décimal. Cette fonction n'étant plus familière aux élèves actuels, c'est pourquoi nous avons utilisé la fonction logarithme népérien.*

Il existe aussi du papier Gausso-arithmétique qui permet de voir si un phénomène peut être modélisé par la loi de Gauss. Nous allons en donner un exemple :

---

[2] On trouvera les graphiques en annexe 2.

[3] Voir graphique en annexe 2.

Quatrième exemple : modèle de la loi de Gauss

Lors d'une fabrication en série d'une pièce, on a mesuré une certaine cote x sur 50 de ces pièces et on a obtenu les résultats suivants (les mesures, effectuées au pied à coulisse, en centimètre, ont été regroupées en classes) :

cote x (en cm)	[3,93 ;3,95[	[3,95 ;3,97[	[3,97 ;3,99[	[3,99 ;4,01[
nombre de pièces	1	2	7	10
fréquence	0,02	0,04	0,14	0,20
fréquence cumulée	0,02	0,06	0,20	0,40

cote x (en cm)	[4,01 ;4,03[	[4,03 ;4,05[	[4,05 ;4,07[	[4,07 ;4,09[	[4,09 ;4,11[
nombre de pièces	16	7	5	1	1
fréquence	0,32	0,14	0,10	0,02	0,02
fréquence cumulée	0,72	0,86	0,96	0,98	1

On reporte x en abscisse (en général on prend la borne supérieure de l'intervalle ou le centre de la classe, cela dépend des conventions) et la fréquence cumulée (en pourcentage) en ordonnée[4]. Si les points sont alignés, on peut modéliser par la loi de Gauss. On trace alors, comme dans les exemples précédents, la droite passant au plus près des points, cette droite s'appelle la droite de Henry[5].

On peut ensuite lire les caractéristiques de la série telles que :

la moyenne m : c'est l'abscisse du point d'intersection de la droite avec la fréquence 0,5 (50%)

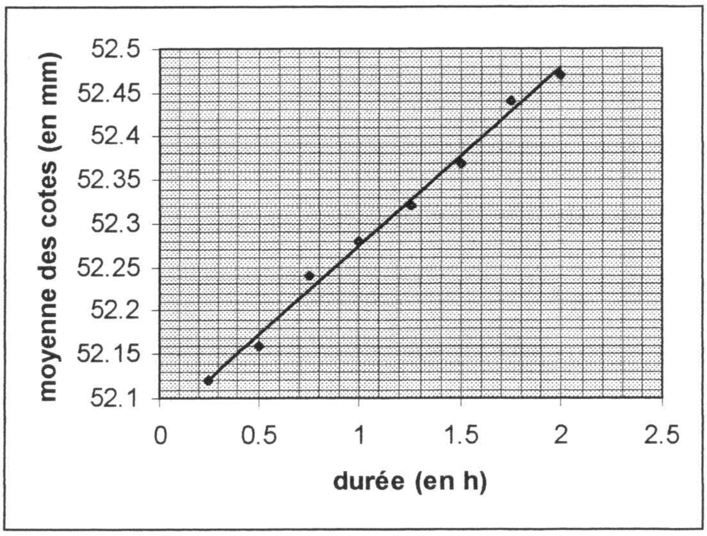
l'écart-type σ : l'abscisse du point d'intersection de la droite avec la fréquence 84,1% donne $m + \sigma$.

[4] Voir graphique en annexe 2.

[5] La méthode apparaît, en 1894, dans le cours du commandant Henry à l'Ecole d'Application d'Artillerie de Fontainebleau.

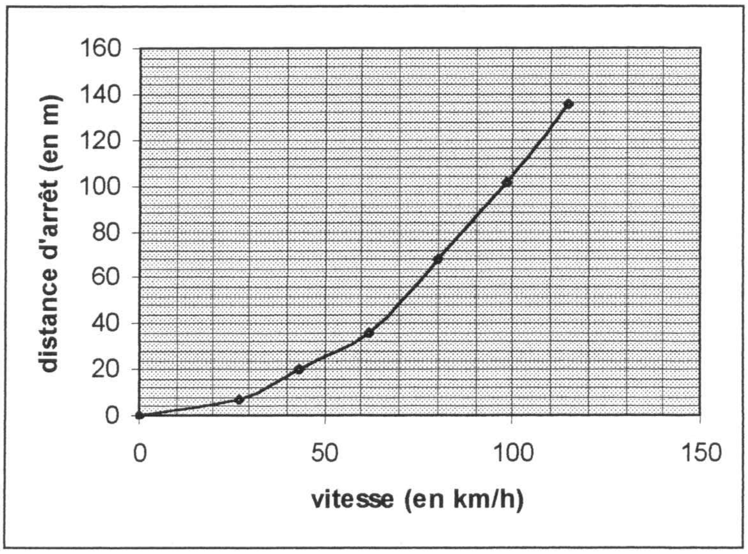
ANNEXE 2

Graphique de l'exemple 1 (modèle affine)

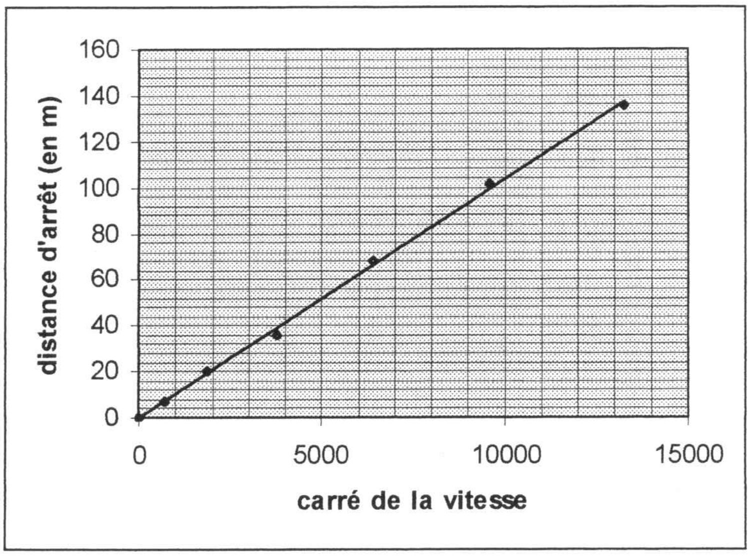


Graphiques de l'exemple 2 (distance d'arrêt)

Premier graphique : distance d'arrêt en fonction de la vitesse :

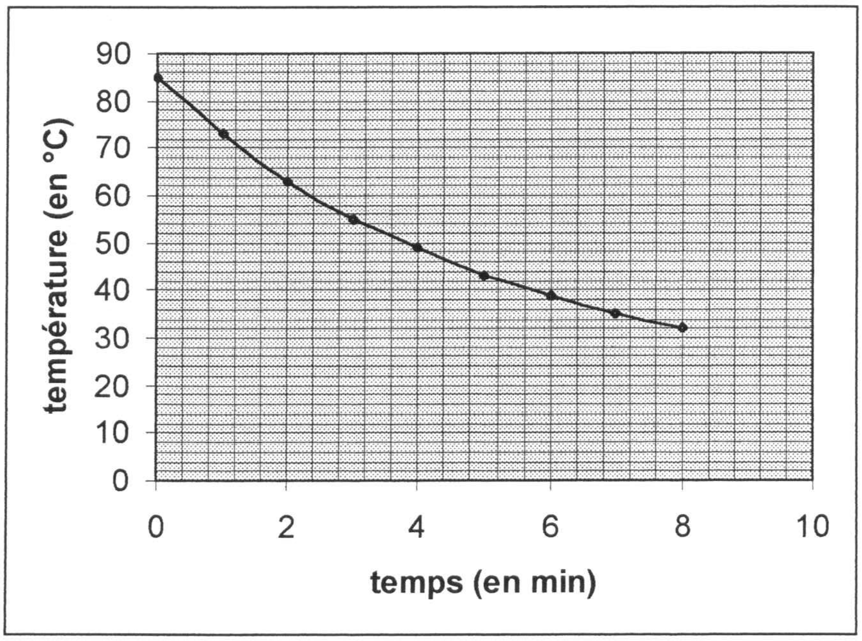


Deuxième graphique : distance d'arrêt en fonction du carré de la vitesse :

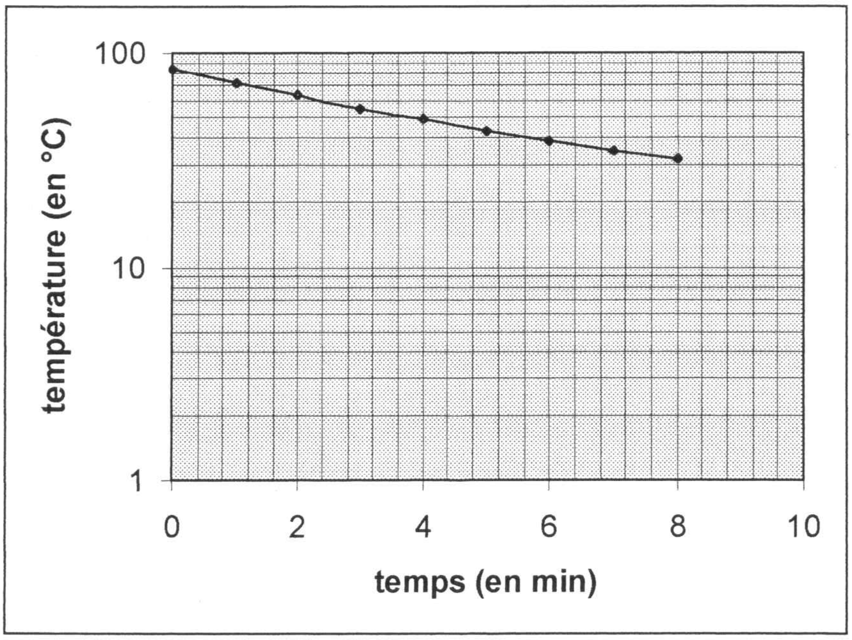


Graphique de l'exemple 3 (modèle exponentiel)

Premier graphique : température en fonction du temps :

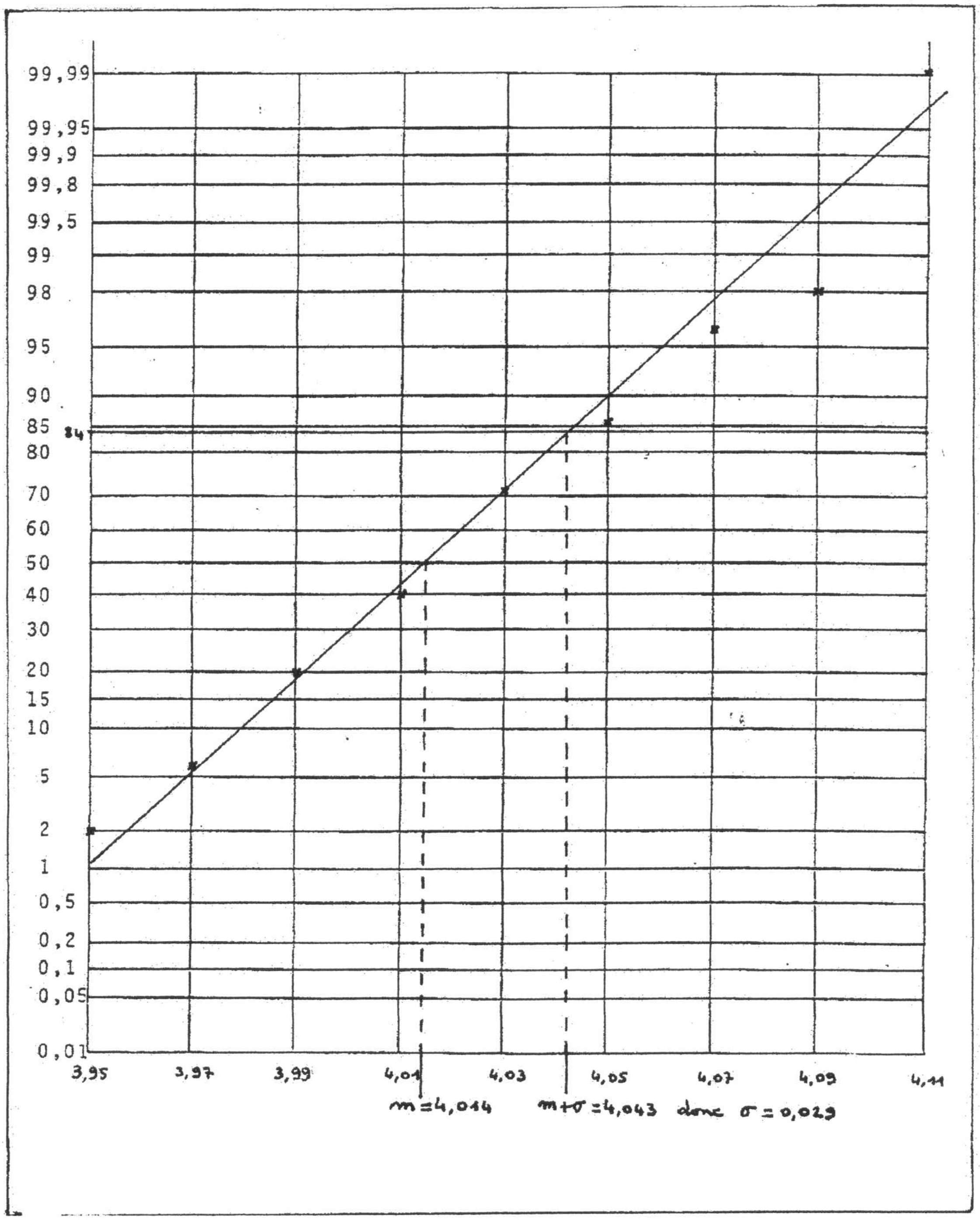


Deuxième graphique : avec une échelle logarithmique en ordonnée :



Graphique de l'exemple 4 (loi de Gauss)

Ce graphique est fait sur du papier Gausso-arithmétique. La droite qui est tracée est la droite de Henry.


99,99
99,95
99,9
99,8
99,5
99
98
95
90
85
84
80
70
60
50
40
30
20
15
10
5
2
1
0,5
0,2
0,1
0,05
0,01
3,95
3,97
3,99
4,01
4,03
4,05
4,07
4,09
4,11
m = 4,014
m+σ = 4,043 donc σ = 0,029

BIOGRAPHIES

BERNOULLI JACQUES (1654-1705) Issu d'une famille flamande émigrée à Bâle à la fin du XVIème siècle, il est le premier d'une famille célèbre en histoire des sciences (voir la généalogie des BERNOULLI). Après avoir voyagé en Europe et rencontré nombre de savants de l'époque, il est nommé professeur à l'université de Bâle en 1687. Sa correspondance avec LEIBNIZ lui permet de développer le calcul infinitésimal mis en place par ce dernier. Mais l'intérêt qu'il porte au calcul des probabilités dès le début des années 1680 conduit à sa notoriété actuelle, mais l'ouvrage qui témoigne de son intérêt pour ce domaine,l'« Ars Conjectandi », ne paraîtra qu'en 1713, l'année de sa mort étant 1705. Il y démontre entre autres, la loi faible des grands nombres.

GALTON FRANCIS (1822-1911), né à Birmingham en Angleterre, est le cousin de Charles Darwin. Touche-à-tout génial et polyvalent, il est avant tout géographe et explorateur. La lecture de " L'origine des espèces " que son cousin a fait paraître en 1859 marque un tournant dans sa vie et il commence à s'intéresser aux méthodes statistiques et à leurs applications à toutes sortes de domaines : génétique, anthropométrie, éducation, psychologie. Il est considéré comme le père de la biométrie et le fondateur de l'eugénisme. Il invente le concept de corrélation et sa mesure par le coefficient de corrélation ; signalons seulement que ce coefficient sert dans la recherche d'une relation de régression, d'où sa notation par la lettre r.

GAUSS KARL–FRIEDRICH (1777-1855) De famille modeste, il bénéficie, au vu de ses aptitudes, d'une bourse qui lui permettra d'étudier jusqu'à intégrer l'université de Göttingen qu'il ne quittera plus jusqu'à sa mort. Si sa thèse de 1800 porte sur le théorème fondamental de l'algèbre, il est aussi célèbre pour ses travaux sur les nombres. Chargé de cartographie, la recherche des coordonnées géographiques de certaines villes au moyen de mesures astronomiques le conduit à utiliser la méthode des moindres carrés et la loi des erreurs qui porte son nom.

GRAUNT JOHN (1620-1674) est né et mort à Londres. Fils d'un marchand drapier, il n'a pas fait d'études ; il a été placé en apprentissage chez un mercier de 1636 à 1641, puis a exercé ce métier jusqu'à sa mort. Autodidacte, John GRAUNT fut un homme en vue et d'une certaine influence à Londres ; il a occupé diverses fonctions publiques entre 1658 et 1674, en particulier il a été conseiller municipal pendant deux ans. Il a été reçu à la *Royal Society* en 1663. Un de ses amis dit de lui qu'il était *« un homme aimé de tous, un ami fidèle. Souvent choisi pour arbitre pour sa prudence et sa droiture, il fut un grand conciliateur. Il avait l'esprit alerte, il était très facétieux et de conversation agréable. Pour lui rendre justice, il était inventif et studieux, aimé de tous, et il se levait tôt chaque matin pour étudier avant l'ouverture de la boutique. Il entendait le Latin et le Français. »*

John GRAUNT est passé à la postérité pour le seul ouvrage qu'il ait publié, ses *Observations sur les tables de mortalité (1662).* Ce travail établit pour la première fois la constance de certains phénomènes biologiques comme, par exemple, le fait qu'il naisse plus de filles que de garçons, que les femmes vivent plus longtemps que les hommes, que la mortalité infantile soit importante.

HUYGENS CHRISTIAAN (1629-1695) Flamand né à La Haye. Son père, CONSTANTIN HUYGENS, correspondant du père Mersenne et ami de Descartes, tient à lui donner, ainsi qu'à ses trois frères une éducation solide qui permet de manifester des dispositions particulières pour le domaine scientifique, où il acquiert une renommée certaine. A l'invitation de LOUIS XIV, il s'installe à Paris, pour faire partie de la toute nouvelle Académie Royale des Sciences. S'il étudie avec succès nombre de courbes « mécaniques », ses travaux concernent aussi les probabilités ; il met ainsi en évidence la notion d' « espérance mathématique ». De retour à La Haye, il meurt en 1695, après avoir quitté la France suite à l'abrogation de l'Edit de Nantes et séjourné à Londres quelques années.

HUYGENS LOUIS (1631-1699) frère du précédent, a bénéficié également de l'éducation scientifique voulue par leur père. Moins brillant que son aîné, il est connu pour la correspondance qu'ils ont entretenue.

LAGRANGE JOSEPH-LOUIS (1736-1813), très précoce, ce fils d'un père français et d'une mère italienne enseigne dès l'âge de 19 ans à l'école d'artillerie de Turin. Exerçant à Berlin, puis à Paris, ses travaux se caractérisent par une approche analytique des problèmes proposés. A la Révolution, il est chargé de la réforme des poids et mesures, et nommé sous l'Empire à l'école Polytechnique.

LAPLACE PIERRE-SIMON (DE) (1749-1827) est né en Normandie. Il quitte sa province en 1767 pour Paris où il se consacre aux mathématiques. Enseignant à l'Ecole Militaire, puis à l'Ecole Polytechnique, il s'intéressera surtout à des problèmes relevant de l'Astronomie. Ceux-ci le conduiront à la théorie des probabilités dont il fera le sujet de deux ouvrages : « Théorie analytique des probabilités » de 1812 et « Essai philosophique sur les probabilités » de 1814 .

LEGENDRE ADRIEN-MARIE (1752-1833) Parisien d'origine, il enseigne à l'Ecole Militaire avant la Révolution . Il met en place, avant même que GAUSS ne l'explicite, la méthode des moindres carrés, pour déterminer la trajectoire de certaines comètes. Il est aussi connu pour ses « Eléments de Géométrie », ouvrage pédagogique qui sera réédité tout au long du XIXème siècle, alors qu'il décède en 1833.

NIGHTINGALE FLORENCE (1820 - 1910) Florence Nightingale, née à Florence, d'où son nom, est plus connue comme infirmière, "la dame à la lampe" de la guerre de Crimée, mettant en œuvre de grandes réformes pour les méthodes sanitaires et l'organisation des hôpitaux, que comme mathématicienne. C'est pourtant l'utilisation et la mise au point de nouvelles méthodes statistiques qui sont à la source de ses idées novatrices.

La carrière de Florence Nightingale est impressionnante, dans cette époque victorienne où la grande majorité des femmes ne suivait aucune étude universitaire et n'avait pas de carrière professionnelle. C'est le père de Florence qui, persuadé que les filles devaient recevoir une bonne éducation, lui fit apprendre l'italien, le latin, le grec, l'histoire et les mathématiques. Elle eut en particulier, comme précepteur de mathématiques, James Sylvester.

En 1854, elle est recrutée, avec 38 autres infirmières pour servir à Scutari pendant la guerre de Crimée. Ses connaissances en statistiques lui permettent alors de mettre en évidence les relations entre la mortalité des militaires et un certain nombre de pratiques médicales. Elle obtient des améliorations sanitaires très nettes grâce à ses méthodes.

Pour faire "parler" ses statistiques, elle utilise des diagrammes, ce qui, à l'époque, est très nouveau. Elle invente en particulier le diagramme polaire.

Elle devient très célèbre et sera consultante sur la santé des armées pendant la guerre civile américaine. Elle sera aussi chargée d'étudier l'amélioration des transports médicaux à travers les grandes étendues canadiennes.

En 1858, elle devient membre de la Royal Statistical Society et en 1874 membre de l'American Statistical Association.

Karl Pearson la considérait comme une prophétesse dans le développement des statistiques appliquées.

Notons aussi qu'elle a consacré une partie de son temps à plaider en faveur de l'éducation des filles, en particulier en mathématiques ; elle a écrit quelques textes pédagogiques, utilisant là encore des méthodes novatrices.

Elle fut pionnière dans un autre domaine de l'enseignement. Sa correspondance avec Quételet montre qu'elle jugeait très important d'introduire l'enseignement des statistiques dans le cursus universitaire.

PEARSON KARL (1857-1936) est né et mort à Londres. Après avoir suivi des études de mathématiques à l'Université de Cambridge et obtenu son diplôme en 1879, il se tourne vers le droit, s'inscrit au Barreau en 1881 puis voyage en Allemagne. Karl PEARSON est un intellectuel qui s'intéresse non seulement aux mathématiques et au droit, mais aussi à la physique, l'histoire et la philosophie. Néanmoins c'est au University College de Londres, où il obtient la chaire de Mathématiques appliquées en 1884, qu'il exercera la majeure partie de sa carrière.

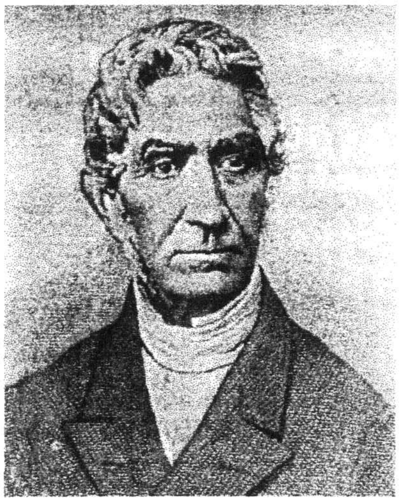
Son admiration pour Galton et son amitié pour Weldon, qui utilise des méthodes statistiques élémentaires pour aborder la théorie de Darwin sur la sélection naturelle, l'amènent à s'intéresser plus particulièrement à la statistique. Il marquera fortement cette discipline. On lui doit en particulier l'étude du coefficient de corrélation, le test du « khi-deux », l'étude de la dispersion : c'est lui qui introduit la notation σ et l'expression *standard déviation* en 1894. Il a aussi construit des tables statistiques de façon à fournir un outil de travail aux utilisateurs de cette discipline.

POISSON SIMEON-DENIS (1781-1840) est né à Pithiviers, mais rejoint Paris assez vite ; admis à dix-huit ans à l'Ecole Polytechnique, il est l'élève de LAPLACE et de LAGRANGE Si ses premiers travaux concernent la géométrie et la mécanique, il se tourne surtout vers la physique mathématique et les probabilités, en particulier par la publication en 1837 de ses « Recherches sur la probabilité des jugements ». C'est dans cet ouvrage que se trouve la loi qui porte son nom, obtenue par passage à la limite de la loi de BERNOULLI.

QUETELET LAMBERT-ADOLPHE-JACQUES (1796-1874) est né à Gand, en Belgique. Devenu mathématicien un peu par hasard, il a l'idée, en 1823 de fonder un observatoire ; afin de s'initier à l'astronomie il rencontre Fourier, Laplace , Poisson à l'Observatoire de Paris et découvre à cette occasion les statistiques. En 1832 il s'installe à l'Observatoire de Bruxelles où il restera jusqu'en 1856. Parallèlement à ses travaux en astronomie et climatologie il s'intéresse de plus en plus à la statistique. Il est l'initiateur du premier Congrès International de Statistique à Bruxelles en 1853.

QUETELET utilise les statistiques pour étudier « l'homme » aussi bien d'un point de vue physique (taille, poids, ...) que d'un point de vue intellectuel et moral (penchant au crime, au mariage, ...). Il défend le principe d'une statistique scientifique s'appuyant sur le calcul des probabilités. Il donne une place prépondérante à la moyenne ; il fait la différence entre la moyenne « objective », qui correspond à quelque chose de réel, et la moyenne arithmétique. Ainsi, avec lui, les statistiques se sont orientées dans une voie nouvelle : les statistiques morales, dont la postérité retiendra particulièrement l'invention de « l'homme moyen ». Cette théorie de « l'homme moyen », qui annonce par exemple la normalisation des qualités, de même que l'utilisation à outrance des mathématiques pour justifier des prises de décision « sociales », ont provoqué un déferlement de critiques. On doit cependant reconnaître que QUETELET a su donner à la statistique du 19ème siècle une impulsion dont l'influence ultérieure est indéniable.

Simeon Denis Poisson



Adolphe Quételet

LES BERNOULLI

- NICOLAS (1623-1708) Apothicaire, Conseiller à la ville de Bâle
  - JACQUES (Jacob) (1654-1705) Professeur à Bâle
    - NICOLAS
  - NICOLAS (1622-1716) peintre
    - NICOLAS (1687-1759) Professeur A Padoue et Bâle
  - JEAN (Johann) (1667-1748) Professeur à Groningue
    - NICOLAS (1696-1726) Berlin et Saint-Petersbourg
    - DANIEL (1700-1782) Physicien Saint-Petersbourg
    - JEAN (1710-1790) Professeur à Bâle
      - JEAN (1744-1807) Directeur de l'Observatoire deBerlin
      - DANIEL (1750-1834)
        - JEAN-CHRISTOPHE (1782-1863) Professeur à Bâle
      - JACQUES (11759-1789) Académicien à Saint-Petersbourg

Bibliographie

Bernoulli J., *Ars conjectandi*, édition bilingue, trad. Norbert Meusnier, IREM de Rouen, 1987.

Bibby J., *Notes towards a history of teaching statistics*, John Bibby, Edimbourg, 1986.

Boyé A., *Des difficultés d'enseigner le hasard et les probabilités*, IREM des Pays de la Loire, Nantes, 1996.
Boyé A., Comairas M. C., *Moyenne, médiane, écart type, quelques regards sur l'histoire pour éclairer l'enseignement des statistiques au lycée*, in Repères IREM, n°48, 2002.

Boyé A., Lefort X., *Mesurer aussi bien la terre que le ciel*, IREM des Pays de la Loire, 1985.

Dacunha-Castelle D., *Chemins de l'aléatoire*, coll. Champs, Flammarion, 1996.

Droesbeke J, Tassi P., *Histoire de la statistique*, Que sais-je ? n° 2527, PUF, Paris, 1997.

Ewald F., *Moyenne et perfection-L'état providence*, in Actes du colloque inter-IREM, Les mathématiques dans la culture d'une époque, Université Louis Pasteur, Strasbourg, 1997.

Feldman J., Labneau G., Matalan B., (sous la direction de*), Moyenne, milieu, centre, histoires et usages*, Ecole des hautes études en sciences sociales, Paris, 1991.

Gauss C. F., *Méthode des moindres carrés*, traduit par J. Bertrand, Mallet-Bachelier, Paris, 1855, Réed. IREM Paris VII, 1996.

Graunt J., *Natural and political observations upon the bills of mortality of the city of London,* 1662, trad. Vilquin E., INED, 1977.

Huygens C., *Correspondance*, in Œuvres complètes, tome VI, Martinus Nijhoff, La Haye, 1920.

Jozeau M. F., *Géodésie au XIX° siècle : de l'hégémonie française à l'hégémonie allemande*, thèse de doctorat, Université Denis Diderot Paris VII, 1997.

Lagrange J. L., *Mémoire sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations, dans lequel on examine les avantages de cette méthode par le calcul des probabilités, et où l'on résout différents problèmes relatifs à cette matière*, in Miscellanea Taurinensia, t.V, 1770-1773, Turin.

Musil R., *L'homme sans qualité*, Points Seuil, Paris, 1996.

Quételet A., *Lettres à son Altesse Royale le Duc de Saxe Cobourg et Gotha sur la théorie des probabilités*, Bruxelles, 1846.

Quételet A., *Sur l'homme et sur le développement de ses facultés ou Essai de physique sociale*, Bachelier, Paris, 1835.

Small H., *Stats and Lamps*, Research conference, Florence Nightingale Museum, Londres, 1998.

Stewart I., *Dieu joue-t-il aux dés ?* coll. Champs, Flammarion, Paris, 1994.

Trystram F., *Le procès des étoiles*, Seghers, Paris, 1979.

Eléments de corrigés des questions et exercices proposés

Chapitre 1 :

Exercice "Buffon"

Soit f la fréquence d'apparition de pile au cours des 4040 lancers et p la probabilité d'obtenir pile.

D'après le théorème de Bernoulli, $P(|f-p| \geq \varepsilon) \leq \dfrac{1}{4 \times 4040 \times \varepsilon^2}$ ou, ce qui est équivalent

$$P(|f-p| < \varepsilon) > 1 - \frac{1}{4 \times 4040 \times \varepsilon^2}$$

. Si l'on veut que cette probabilité soit supérieure à 0,95 il suffit que, $1 - \dfrac{1}{4 \times 4040 \times \varepsilon^2} \geq 0{,}95$

ce qui donne : $\varepsilon \geq 0{,}035$

La probabilité p appartient donc à l'intervalle [0,472 ; 0,542] avec une probabilité au moins égale à 0,95. Or p vaut 0,5, qui appartient bien à cet intervalle. On peut donc raisonnablement penser que la pièce de Buffon était équilibrée.

Chapitre 2

Question 1 :

a) Si les deux résultats sont ρ et ρ - 1, le milieu est $\rho - \dfrac{1}{2}$ et l'erreur commise en prenant ce milieu est $-\dfrac{1}{2}$

b)

Assemblages de résultats possibles	ρ et ρ - 1	ρ et ρ + 1	ρ et ρ	ρ - 1 et ρ - 1	ρ - 1 et ρ + 1	ρ + 1 et ρ + 1
milieux	ρ - 1/2	ρ + 1/2	ρ	ρ - 1	ρ	ρ + 1
erreurs	- 1/2	+ 1/2	0	- 1	0	+ 1

Question 2 :

Assemblages de résultats possibles	ρ ; ρ ; ρ	ρ ; ρ ; ρ - 1	ρ ; ρ ; ρ + 1	ρ ; ρ - 1 ρ - 1	ρ ; ρ - 1 ρ + 1	ρ ; ρ + 1 ρ + 1	ρ - 1 ; ρ - 1 ; ρ - 1	ρ - 1 ; ρ - 1 ; ρ + 1	ρ - 1 ; ρ + 1 ; ρ + 1	ρ + 1 ; ρ + 1 ; ρ + 1
milieux	ρ	ρ - 1/3	ρ + 1/3	ρ - 2/3	ρ	ρ + 2/3	ρ - 1	ρ - 1/3	ρ + 1/3	ρ + 1
erreurs	0	- 1/3	+ 1/3	- 2/3	0	2/3	-1	- 1/3	+ 1/3	+1

Question 3 :

a) Notons e l'erreur

$\|e\|$	0	1/2	1
Probabilité de $\|e\|$	2/6	2/6	2/6

La probabilité que $|e| \leq 1/2$ est de 4/6 ou 2/3 ; la probabilité que $|e| \leq 1/3$ est de 2/6 ou 1/3.

b)

$\|e\|$	0	1/3	2/3	1
Probabilité de $\|e\|$	2/10	4/10	2/10	2/10

La probabilité que $|e| \leq 1/2$ est de 6/10 ; la probabilité que $|e| \leq 1/3$ est de 6/10.

Question 5 :

L'univers est constitué des n-uplets constitués de chiffres 0, 1 et -1.

Le nombre d'éléments de l'univers est $(a + 2b)^n$.

Question 6 :

a) Lors du lancer de 3 dés il y a 7^3 possibilités, soit 343.

Pour obtenir 3, il faut les faces 1, 1, 1 et - 3 les faces -1 , -1 , -1. Sur chaque dé, il y a trois faces de chaque type. Donc le nombre de façons d'obtenir 3 ou -3 est 54.

Pour obtenir 2, il faut deux faces 1 et une face 0. En agençant de toutes les façons possibles la face 0 parmi les faces 1, et en sachant qu'il y a sur chaque dé une face 0 mais trois faces 1, cela donne 27 possibilités pour 2 et autant pour - 2. Donc au total, 54 possibilités d'obtenir 2 ou -2.

Pour obtenir 1, il faut deux faces 0 et une face 1ou deux faces 1 et une face - 1 ; en déterminant toutes les possibilités d'agencer ces faces les unes par rapport aux autres, et en tenant compte du fait que sur chaque dé il y a trois faces 1, trois faces - 1 et une face 0, nous obtenons 90 façons d'avoir 1 et autant d'avoir - 1, donc 180 façons d'avoir 1 ou - 1.

Pour obtenir 0, il faut trois faces 0 ou une face 0 une face 1 et une face - 1.

Nous obtiendrons 55 possibilités en tout.

Nous vérifions que 54 + 54 + 180 + 55 = 343

	0	±1	± 2	± 3
probabilités	55/343	180/343	54/343	54/343

La probabilité que $|e| = 0$ est de 55/343 ; que $|e| \leq 1/3$ est de 235/343 ; et que $|e| \leq 2/3$ est de 289/343.

b) S'il y a 4 dés, le nombre total de possibilités est de 7^4, soit 2401.

En procédant comme précédemment, c'est-à-dire en déterminant toutes les possibilités d'obtenir une somme donnée et en tenant compte du fait que sur chaque dé il y a une face 0 et trois faces 1 et trois faces - 1, nous obtiendrons le tableau suivant :

Nombre de points	± 4	± 3	± 2	± 1	0
Nombre de possibilités	162	216	756	672	595
probabilités	162/2401	216/2401	756/2401	672/2401	595/2401

La probabilité que $|e| = 0$ est de 595/2401 ; que $|e| \leq 1/4$ est de 1267/2401 ; que $|e| \leq 1/2$ est de 2043/2401 ; que $|e| \leq 3/4$ est de 2259/2401.

c) Si on lance 5 dés, le nombre total de possibilités est de 7^5 soit 16807 cas.

Nombre de points	± 5	± 4	± 3	± 2	± 1	0
Nombre de possibilités	486	810	2970	3420	6510	2611
probabilités	486/16807	810/16807	2970/16807	3420/16807	6510/16807	2611/16807

La probabilité que $|e| = 0$ est de 2611/16807 ; que $|e| \leq 1/5$ est de 9121/16807 ; que $|e| \leq 2/5$ est de 12541/16807 ; que $|e| \leq 3/5$ est de 15511/16807 ; que $|e|$ 4/5 est de 16321/16807.

d)

Probabilité que l'erreur commise en prenant le milieu des observations ne surpasse pas, en valeur absolue :

Valeurs du nombre n des observations	0	1/n	2/n	3/n	4/n
3	55/343	235/343	289/343	....	....
4	595/2401	1267/2401	2043/2401	2259/2401	....
5	2611/16807	9121/16807	12541/16807	15511/16807	16321/16807

Question 7 :

$(1 + 3(x + x^{-1}))^3 = 55 + 90\,x + 90\,x^{-1} + 27\,x^2 + 27\,x^{-2} + 27\,x^3 + 27\,x^{-3}$.

Nous retrouvons bien les dénombrements du cas de 3 dés.

$(1 + 3(x + x^{-1}))^4 = 595 + 336x + 336x^{-1} + 378x^2 + 378x^{-2} + 108x^3 + 108x^{-3} + 81x^4 + 81x^{-4}$.

Ici aussi nous retrouvons bien les dénombrements du cas de 4 dés.

Question 8 : il s'agit de la moyenne pondérée entre les nombres p, q, r … affectés respectivement des coefficients a, b, c, ….

Question 9 :on aurait envie de parler de points sur une droite, de leurs **abscisses** sur la **droite des réels**, de **vecteurs**, de **barycentre** …

Chapitre 4

1. Louis a fait une **interpolation linéaire**.

a. Une personne de 36 ans vit jusqu'à 53,5 ans

38 ans vit jusqu'à x ans

46 ans vit jusqu'à 61 ans

donc : $\dfrac{x-53,5}{38-36}=\dfrac{61-53,5}{46-36}$ d'où $x=53,5+\dfrac{7,5}{10}\times 2=53,5+1,5=55$

b. Réciproquement, soit x l'âge actuel d'une personne vivant jusqu'à 56,5 ans.

On a donc : $\dfrac{x-36}{56,5-53,5}=\dfrac{46-36}{61-53,5}$ d'où $x=36+\dfrac{10}{7,5}\times 3=40$.

Christian a donc 40 ans en 1669.

2. En appliquant la méthode graphique donnée par Christian, on trouve que Louis vivra jusqu'à 52 ans environ (un peu moins) et Christian jusqu'à 53 ans.

Chapitre 5 :

Tous les secteurs ayant le même angle au centre (360°/12 = 30°), leur aire est proportionnelle au carré de leur rayon.

Mois	Nombre de décès dus aux maladies	Nombre total de décès	Pourcentage
Janvier 1855	3228	3631	89%
Février 1855	2340	2770	84%
Mars 1855	1378	1594	86%

Table des matières