

# ÉMERGENCE HISTORIQUE DES LOIS À DENSITÉ

## DES PISTES POUR L'ENSEIGNEMENT

Charlotte DEROUET<sup>1</sup>

### TITLE

The historical emergence of density distributions. Hints for teaching

### RÉSUMÉ

Depuis 2012, l'enseignement des lois à densité fait partie de l'enseignement des probabilités de la classe de terminale de nombreuses filières. Cet article a pour objectif de présenter un aperçu de la naissance historique de la notion de fonction de densité de probabilité dans le but d'en dégager quelques éléments qui semblent intéressants à prendre en compte lors de l'introduction de cette notion en classe. Il s'agit de pointer plusieurs phases de la construction historique des lois à densité en mettant en évidence le rôle joué par la statistique. Cette étude historique et épistémologique montre notamment la place de la statistique dans l'émergence de la notion comme point de départ d'un problème concret : les erreurs de mesures observées. Dans cet article, nous mettrons en évidence des éléments didactiques qui peuvent être exploités pour introduire la notion de fonction de densité en classe de terminale scientifique ou autres filières, mais aussi dans l'enseignement supérieur.

*Mots-clés : fonction de densité de probabilité, probabilités, statistique, enseignement secondaire, didactique des mathématiques.*

### ABSTRACT

Since 2012, the teaching of density distributions is part of the teaching of probability during the last year of French highschool of many tracks. This article aims at presenting an overview of the historical birth of the notion of density function in order to identify elements to be taken into account to introduce this notion in classes. Several phases of the historical construction of density distributions are identified. This historical and epistemological study shows the role of statistics in the emergence of the notion, especially continuous distributions are the answer to a concrete problem : the errors of observed measurement. From this study, some didactic elements are drawn that can be exploited to introduce the notion of density function at the end of highschool in the scientific track or others, but also in higher education.

*Keywords: density function, probability, statistics, secondary school, mathematics education.*

## 1 Introduction

À partir de 2002, la loi uniforme sur  $[0; 1]$  et les lois exponentielles ont été introduites dans le programme de la classe de terminale S, mais cette partie du programme n'était alors

---

<sup>1</sup>ESPE de l'Académie de Paris (Université Paris Sorbonne) & Laboratoire de Didactique André Revuz (Université Paris Diderot), charlotte.derouet@espe.unistra.fr

que peu développée. Depuis la rentrée 2012, l'enseignement des lois à densité (en particulier la loi normale) en terminale s'est généralisé dans de nombreuses filières générales et technologiques. Si l'on prend le cas de la classe de terminale scientifique, il est maintenant demandé d'aborder à ce niveau scolaire les lois uniformes, exponentielles et normales. Les différentes lois sont définies par leur fonction de densité associée. Dans notre thèse de doctorat (Derouet, 2016), nous nous posons la question suivante : comment introduire la notion de fonction de densité de probabilité en classe de terminale S ?

Dans cet article, nous présenterons des éléments historiques et épistémologiques sur la naissance des lois à densité qui mettent notamment en évidence la place de la statistique dans la genèse de la notion de fonction de densité. Nous en extrairons ensuite des éléments didactiques qui semblent pertinents à prendre en compte pour introduire la notion de fonction de densité en classe de terminale. Ces pistes, associées à d'autres issues d'analyses du programme, des documents *Ressources* et des manuels (Derouet & Parzys, 2016), ont permis la conception de séances d'introduction qui ont été expérimentées en classe de terminale S dans le cadre de notre thèse (Derouet, 2016). Nous ne présenterons pas ces résultats dans cet article.

Notre point de vue sur la genèse des lois à densité est orienté par le fait que nous sommes didacticienne et non historienne. Nous avons retenu les éléments historiques, autour de la théorie des erreurs, qui nous paraissent révélateurs de l'émergence de la notion de fonction de densité et nous mettrons en avant les liens qui existent avec la statistique et qui ont contribué à cette construction épistémologique. Bien que non exhaustive, l'étude historique et épistémologique est, comme l'écrit Dorier (2000), « *une phase fondamentale pour que le chercheur puisse prendre ses distances par rapport aux enjeux didactiques* » (p. 9). Plus spécifiquement, Robinet (1984) insiste sur le fait que :

*Les concepts doivent être introduits par leur fonctionnement et par conséquent « au plus près de leur sens » comme outil, donc dans les problèmes où ils sont une réponse. Historiquement les problèmes, pour lesquels ils ont été inventés, remplissent ces conditions, mais d'une part ils ne sont pas toujours adaptables à l'enseignement, d'autre part le fonctionnement du concept le fait apparaître parfois « après coup » comme plus nécessaire dans d'autres types de problèmes.*  
(p. 18)

Pour cette étude historique et épistémologique, nous nous sommes appuyé sur des articles de Parzys (2013), Bru (2006) et Pichard (2005). Ces différentes sources portent sur des questions historiques liées à la genèse de la loi normale avec un regard didactique. Bien que ces articles soient centrés sur la loi normale, il est possible d'en extraire des informations plus générales sur les lois à densité. Nous avons complété notre étude avec deux livres portant sur l'histoire de la statistique (Stigler, 1986 ; Drosbeke & Tassi, 1990). Enfin, nous avons aussi étudié quelques textes originaux que nous citerons en temps voulu.

Cet article se compose de deux parties. Dans la première partie, la plus développée, nous détaillerons plusieurs phases qui nous paraissent importantes dans l'émergence de la notion de fonction de densité à travers l'histoire de la théorie des erreurs. Dans la seconde partie, nous dégagerons des pistes (non exhaustives) pour l'enseignement qui découlent de cette étude.

## 2 Quelques éléments historiques. La théorie des erreurs : naissance des lois à densité

Bien que le problème de Bernoulli, qui débouche plus tard sur le théorème que nous appelons maintenant théorème de Moivre Laplace (cas particulier du théorème limite central pour les lois binomiales), fasse apparaître une intégrale du type gaussien, ce n'est pas ce problème qui est à l'origine des lois à densité. Les premières distributions à densité sont en effet apparues avec la naissance de la théorie des erreurs. Au cours du siècle des Lumières, l'astronomie et la géodésie<sup>2</sup> sont en plein essor. Les instruments de mesures sont de plus en plus précis mais se pose alors la question des écarts entre les différentes mesures (observées) d'une même quantité. Armatte (2004) détaille les enjeux de cette démarche. Les mathématiciens cherchent alors à déterminer la « vraie » ou sinon la « meilleure » valeur de la quantité mesurée et à trouver une « loi des erreurs ». La théorie des erreurs va alors naître : il s'agit de la théorie mathématique des erreurs d'observation qui permet, par exemple, à l'astronomie de précision de pouvoir concilier les résultats des observations avec la mécanique céleste (les résultats théoriques) qui en rend compte.

Lorsque l'on fait une mesure d'une quantité à l'aide d'un instrument (optique, par exemple), même si on enlève les erreurs systématiques de l'instrument, il reste les erreurs inévitables : défauts d'attention, variations anarchiques de la réfraction atmosphérique, vibrations diverses, etc. Toutes ces causes accidentelles d'erreurs ont des conséquences sur la précision de la mesure. Si l'on mesure plusieurs fois la même quantité, on trouve à chaque fois des valeurs différentes et cela est d'autant plus vérifié lorsque les observations sont indirectes (par exemple la masse de Jupiter) car elles ne donnent le résultat cherché que par l'intermédiaire d'équations mettant en jeu plusieurs mesures de natures différentes (Bru, 2006).

La théorie des erreurs cherche à modéliser la dispersion des erreurs de mesure d'une valeur qui doit être centrale et qui devrait être la valeur théorique. Dans son principe, elle consiste à décrire la loi de probabilité des erreurs (écarts entre la « vraie » valeur d'une quantité et les mesures fournissant des valeurs observées). Un des précurseurs dans ce domaine est Galilée (1564-1643) qui, en 1632, s'intéresse à la détermination de la distance entre la Terre et une nouvelle étoile. Il dispose de 78 mesures d'élévation de cette étoile, fournies par 13 observateurs répartis à plusieurs endroits de la planète. Dans son écrit intitulé *Dialogo sopra i due massimi sistemi del mondo : Tolemaico e Copernicano* (1632), Galilée écrit :

*Étant donné que ces observateurs sont capables, qu'ils ont tous fait des erreurs, et que leurs erreurs doivent être corrigées pour nous permettre d'avoir la meilleure information possible à partir de ces observations, nous nous consacrerons à appliquer les modifications minimales et les corrections les plus petites possibles, juste assez pour ôter les observations de l'impossible et les remettre dans le possible... (cité dans Droesbeke & Tassi, 1990, p. 11)*

---

<sup>2</sup>La géodésie est la science destinée, à l'origine, à la mesure de la Terre et au tracé des cartes.

## 2.1 Une approche purement théorique

En 1757, Simpson (1710-1761) explique que ce ne sont pas les observations qui importent, mais les écarts entre ces observations et la « vraie » valeur, qui est inconnue. Il fait une analyse probabiliste de ces écarts, qu'il appelle erreurs (d'où le nom de « théorie des erreurs » introduit par Lambert), et suppose que les erreurs positives et négatives également distantes de la valeur théorique ont la même probabilité de survenir (Parzys, 2013). Il suppose donc la symétrie de la distribution. La publication de Simpson est considérée comme la première portant sur les lois des erreurs. Il introduit d'abord la loi uniforme discrète, puis la loi triangulaire discrète des erreurs. À l'aide de la fonction génératrice (utilisée par De Moivre et théorisée par Laplace), Simpson obtient ensuite la loi de probabilité de la somme de  $n$  erreurs indépendantes. Par passage à la limite, il introduit la distribution de la moyenne arithmétique d'erreurs, suivant chacune une loi triangulaire continue définie par la fonction de « facilité » (que nous appelons maintenant fonction de densité) :

$$f(x) = \frac{a - |x|}{a^2}, \text{ avec } -a \leq x \leq a. \quad (1)$$

Le fait d'introduire la moyenne arithmétique est très courant<sup>3</sup> pour réduire les imprécisions des mesures. À propos de la figure 1, Simpson écrit :

*Let, then, the line AB represent the whole extent of the given interval, within which all the observations are supposed to fall; and conceive the same to be divided into an exceeding great number of very small, equal particles, by perpendiculars terminating in the sides AD, BD of an isosceles triangle ABD formed upon the base AB : and let the probability or chance whereby the result of any observation tends to fall within any of these very small intervals Nn, be proportional to the corresponding area NMmn, or to the perpendicular NM ; then, since these chances (or areas) reckoning from the extremes A and B, increase according to the terms of the arithmetical progression 1, 2, 3, 4, &c. it is evident that the case is here the same with that in the latter part of Prop. II [where the discrete triangular distribution was studied]; only, as the number v (expressing the particles in AC or BC) is indefinitely great, all (finite) quantities joined to v, or its multiples, with the signs of addition or subtraction, will here vanish, as being nothing in comparison to v. (Simpson, 1757, cité dans Stigler, 1986, p. 96) <sup>4</sup>*

<sup>3</sup>On sait maintenant que la moyenne arithmétique est le meilleur estimateur de l'espérance d'une loi limite.

<sup>4</sup>Traduction contemporaine par nos soins : *Le segment [AB] représente l'étendue entière de l'intervalle donné, dans lequel toutes les observations sont supposées tomber ; ce segment est divisé en un très grand nombre de très petites parties régulières, par des perpendiculaires délimitées par les côtés [AD], [BD] d'un triangle isocèle ABD de base [AB] : la probabilité ou la chance que le résultat d'une quelconque observation tende à tomber dans l'un des petits intervalles Nn, est proportionnelle à l'aire correspondante de NMmn, ou à la longueur de la perpendiculaire [NM] ; donc, puisque ces chances (ou aires) dépendant des extrémités A et B croissent suivant les termes de la progression arithmétique 1, 2, 3, 4 etc., il est évident que le cas ici est le même que celui de la partie sur la loi triangulaire discrète ; seulement, comme le nombre v (exprimant le nombre de subdivisions sur [AC] ou [BC]) est infiniment grand, toutes les quantités (finies) reliées à v, ou ses multiples, avec un signe positif ou négatif, disparaîtront, n'étant rien comparées à v.*

C. Derouet

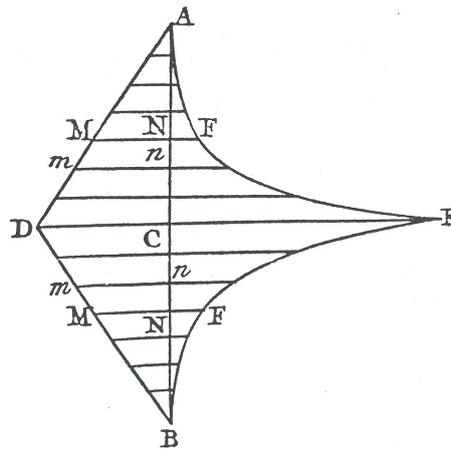


FIGURE 1 – Graphique de la loi d'erreurs triangulaire  $ADB$  et de la loi de la moyenne des erreurs  $AFEFB$  (Simpson, 1757, cité dans Stigler, 1986, p. 96).

Nous voyons apparaître ici une « courbe de densité », bien qu'elle ne soit pas sous la forme actuelle (la théorie des fonctions est encore en évolution à cette époque), et la citation mentionne bien le lien entre probabilité et aire sous la courbe : « la probabilité (ou la chance) que le résultat d'une quelconque observation tende à tomber dans l'un des petits intervalles  $Nn$  est proportionnelle à l'aire correspondante  $NMmn$  »<sup>5</sup>.

En parallèle, en 1757, Boscovic (1711-1787) écrit que la meilleure estimation est celle pour laquelle :

- la somme des écarts (erreurs de mesures) positifs est égale à celle des écarts négatifs ;
- la somme des valeurs absolues des écarts est aussi petite que possible.

Dans ce cas, la « vraie » valeur n'est pas la moyenne arithmétique, mais la médiane de la série de mesures. En effet, la somme des écarts des valeurs observées à une valeur  $a$  est minimale lorsque  $a$  est la médiane, alors que la somme des carrés de ces écarts est minimale lorsque  $a$  est la moyenne arithmétique.

En 1760, Lambert (1728-1777) consacre une partie de son traité *Photometria sive de mensura* (1760) à l'étude de courbes de « facilité », continues, symétriques et unimodales. Il introduit la méthode du maximum de vraisemblance (qui ne porte pas encore ce nom) en recherchant la position du centre d'une distribution continue symétrique. On voit apparaître dans ce traité une courbe (cf. figure 2). Lambert explique :

*Soit AC la valeur vraie à déterminer au moyen d'expériences, soit CB et CD les valeurs maximales de part et d'autre. Que les ordonnées QM, PN, RL, SK de la courbe BMLD représente les retours des erreurs CQ, CP, CR, CS, à savoir leurs fréquences. Nous appellerons ici celles-ci fréquences*

<sup>5</sup>Traduit par nos soins.

*Émergence historique des lois à densité. Des pistes pour l'enseignement*

*absolues ou vraies pour les distinguer des fréquences observées. Il est clair que celles-ci coïncident avec celles-là si l'expérience est infiniment répétée. Mais puisque cela n'arrive jamais, il faut découvrir ce qui demeure pour un nombre fini d'expériences (Lambert, 1760, cité dans Lefort & Boyé, 1996, p. 244)*

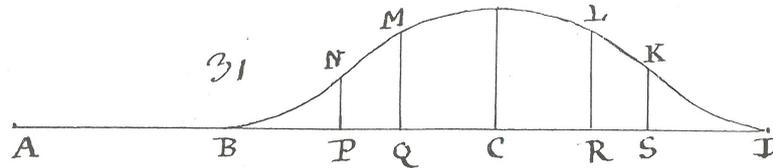


FIGURE 2 – Graphique d'une loi de « facilité » des erreurs (Lambert, 1760, cité dans Lefort & Boyé, 1996).

Il ne s'agit pas exactement d'une courbe de densité. Lambert précise que l'ordonnée correspond à la fréquence. Il s'agit en effet d'un diagramme en bâtons des fréquences « vraies » dont les sommets sont ensuite reliés par une courbe continue. Cet extrait met bien en évidence le lien entre fréquences (observées) et probabilités (fréquences absolues ou vraies), que nous pouvons justifier aujourd'hui par la loi faible des grands nombres.

En 1768, Lagrange (1736-1813) rédige un *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*. Dans ce mémoire, on retrouve les résultats de Simpson, ainsi que de nouvelles lois continues : la loi uniforme sur un segment, la distribution parabolique et la distribution sinusoidale. Jusque là, les mathématiciens considèrent que les erreurs doivent être bornées : une valeur trop éloignée des autres sera éliminée.

Laplace est le premier à considérer une loi de « facilité » sur un domaine infini. En 1774, il indique quelques conditions que doit vérifier la courbe de « facilité » des erreurs (Parzys, 2013) :

*La loi suivant laquelle cette vraisemblance [probabilité d'apparition] diminue à mesure que l'observation s'éloigne de la vérité nous est inconnue. Supposons donc [figure 3] que le point V soit le véritable instant du phénomène [il s'agit d'une mesure de temps] [...] en nommant  $x$  l'abscisse VP, et  $y$  l'ordonnée correspondante PM, nous représenterons l'équation par celle-ci :  $y = \phi(x)$ .*

*Or voici les propriétés de cette courbe [cf. figure 3] :*

1. elle doit être partagée en deux parties entièrement semblables par la droite VR, car il est tout aussi probable que l'observation s'écartera de la vérité à droite comme à gauche ;
2. elle doit avoir pour asymptote la ligne KP, parce que la probabilité que l'observation s'éloigne de la vérité à une distance infinie est nulle ;

C. Derouet

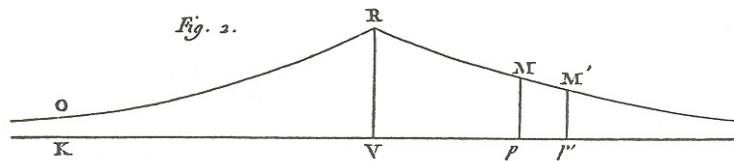


FIGURE 3 – Courbe de la première loi de Laplace (Laplace, 1774, cité dans Stigler, 1986, p. 107)

3. l'aire entière de cette courbe doit être égale à l'unité, puisqu'il est certain que l'observateur tombera sur un des points de la droite  $KP$ .

Laplace fait à nouveau apparaître, après Simpson, le lien entre la probabilité et l'aire sous la courbe. On voit ici émerger des contraintes sur la fonction de densité :

- la courbe de la fonction est symétrique par rapport à la droite verticale de la valeur «réelle»,
- la limite de la fonction en  $+\infty$  est nulle,
- l'aire sous la courbe vaut 1.

On retrouve, dans les deux derniers points, les propriétés qu'une fonction de densité doit vérifier sur  $\mathbb{R}$ , de façon générale. La propriété de la positivité de la fonction reste implicite.

Laplace étend l'intervalle des possibles à  $\mathbb{R}$  en introduisant la loi appelée maintenant la première loi de Laplace, aussi appelée loi double exponentielle définie par la fonction  $f(x) = \frac{k}{2}e^{-k|x|}$ . Dans ses travaux ultérieurs, il rencontre fréquemment l'intégrale  $\int e^{-t^2} dt$ , qui lui semble importante et explique qu'il serait utile d'en faire une table sur  $] -\infty; 0]$  (Parzys, 2013). Mais il faut attendre Gauss (1777-1855), en 1809, pour introduire la loi de Gauss, aujourd'hui appelée loi normale. Elle est introduite en recourant à la méthode des moindres carrés. À la même époque, en 1810, Laplace introduit la «seconde loi des erreurs», qui n'est autre que la loi de Gauss : il montre que la distribution de la moyenne arithmétique de  $n$  erreurs de mesures d'une même quantité, indépendantes, peut être approchée par la distribution d'une loi normale, avec une erreur d'approximation de plus en plus faible lorsque  $n$  tend vers l'infini. En 1824, Poisson considère la loi définie sur  $\mathbb{R}$  par la fonction  $f(x) = \frac{1}{\pi(1+x^2)}$ , que nous appelons aujourd'hui loi de Cauchy.

De la théorie des erreurs émergent les premières lois à densité (toutes symétriques) et la notion de fonction de densité (qui ne porte pas encore ce nom) et notamment ses propriétés. Nous voyons apparaître une relation entre les probabilités et la statistique. La courbe de «facilité» représente alors la fréquence «théorique».

Nous avons donné un rapide aperçu des scientifiques de l'époque ayant travaillé sur la théorie des erreurs de manière théorique. Cependant, les astronomes et géodésiens de terrain ont aussi contribué à l'avancée sur ces lois de probabilités.

## 2.2 Un retour aux données : la vérification empirique

Dans les années 1830, astronomes et géodésiens ont commencé à dresser des tableaux statistiques pour rassembler toutes leurs données, ce qui a permis de confronter les données expérimentales aux modèles théoriques proposés par les mathématiciens. Selon Bru (2006), un des premiers à l'avoir fait serait Bessel (1784-1846). Dans l'annexe 1, nous pouvons voir des comparaisons entre des distributions de valeurs absolues d'erreurs résiduelles de trois groupes d'observations et des distributions théoriques calculées. Bru (*ibid.*) précise que le tableau de fréquences théoriques associé ressemble fortement aux tables de la courbe de Gauss.

Dans le cours de *Calcul des probabilités et Théorie des erreurs* (1852), Liagre présente des données des travaux de Bessel et compare les valeurs d'erreurs observées et les valeurs théoriques (cf. figure 4), à l'aide de la table de l'intégrale  $\frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$  (cf. Annexe 2). Les valeurs sont très proches, ce qui valide le choix de la loi de Gauss comme loi des erreurs. Il en est ainsi car les phénomènes étudiés sont effectivement de type gaussien (il s'agit de la somme de multitudes d'interventions du hasard indépendantes).

À la suite, Liagre justifie par la pratique le fait de considérer une courbe de « facilité » définie sur  $\mathbb{R}$  et non sur un intervalle borné :

*Il arrive fréquemment, comme ici, que l'observation présente plus de grandes erreurs que la théorie n'en prévoit. Cela prouve la légitimité de l'hypothèse que nous avons faite [...] en intégrant le dénominateur entre les limites  $-\infty$  et  $+\infty$ , au lieu de l'intégrer entre des limites finies ; car le contraire aurait lieu si l'inexactitude de notre hypothèse devait avoir quelque influence dans la pratique. (Liagre, 1852)*

Il justifie bien ce choix des bornes infinies par un retour sur les données expérimentales.

Bien que les tableaux de fréquences soient de plus en plus courants, il faut attendre encore un peu pour voir apparaître des courbes de fréquences, qui correspondent à des diagrammes en bâtons. Elles deviendront systématiques seulement à la fin du XIX<sup>e</sup> siècle.

Quetelet (1796-1855) «montre» que la théorie des erreurs peut s'appliquer à d'autres domaines que les erreurs de mesures, et notamment il repère que les tours de poitrine des soldats écossais sont régis par la loi de Gauss. Pour lui, c'est aussi le cas pour différentes mesures de caractéristiques humaines (taille, tour de poitrine, taille de crâne...) et pour les phénomènes en science morale (ou physique sociale). Pour des données d'observation, Quetelet utilise l'ajustement à une courbe des erreurs comme un test sur l'homogénéité de la population. Si l'histogramme construit à partir des observations n'a pas la forme d'une courbe en cloche, cela montre une hétérogénéité (à déterminer). Bertillon (1821-1883) offre un exemple frappant sur les tailles des conscrits du département du Doubs en 1863 avec une courbe à double bosse (cf. figure 5). L'hétérogénéité est expliquée ici par un mélange de types de personnes différents.

Dans son essai pour donner un fondement expérimental à la théorie de l'évolution de Darwin, Galton (1822-1911) utilise la méthodologie de Quetelet, en particulier celle

C. Derouet

■●■. — Comme exemple de l'accord qui règne entre la théorie et l'expérience, nous choisisons une série de 470 observations faites par Bradley, le modèle des observateurs suivant l'expression de Laplace, pour déterminer directement la différence d'ascension droite entre le soleil et l'une des deux étoiles Atair et Procyon. Bessel, qui a discuté ce sujet, a trouvé que l'erreur probable  $r = 0'',2637$ ; d'où  $0'',1 \pm 0,3792 r = 0,3792$ , en prenant  $r$  pour unité. Cherchant ensuite le nombre d'erreurs qui, suivant la théorie, doivent tomber entre  $0'',0$  et  $0'',1$ ;  $0'',2$ ;  $0'',3$ ;  $0'',4$ ;  $0'',5$ ... , il a calculé, au moyen de la table mentionnée au § 99 (table n° 5), la valeur de l'intégrale  $P_2$  pour ces diverses limites \*. On trouve ainsi :

Pour $0'',1$	... $0,3792$	le nombre	$0,20186$ ,
" $0,2$	... $0,7584$	"	$0,39402$ ,
" $0,3$	... $1,1376$	"	$0,55705$ ,
" $0,4$	... $1,5168$	"	$0,69372$ ,
" $0,5$	... $1,8960$	"	$0,79904$ .
" $0,6$	... $2,2752$	"	$0,87511$ .
" $0,7$	... $2,6544$	"	$0,92661$ .
" $0,8$	... $3,0336$	"	$0,95926$ ,
" $0,9$	... $3,4128$	"	$0,97866$ ,
" $1,0$	... $3,7920$	"	$0,98983$ .
	... $\infty$	"	$1,00000$ .

Si l'on retranche chaque nombre du suivant, et qu'on multiplie le reste par le nombre des observations, ou par 470, on trouve

ENTRE	NOMBRE	D'APRÈS	
	DES ERREURS.	la théorie.	l'observation.
$0'',0$ et $0'',1$	0,20186	95	94
$0,1$ $0,2$	0,48916	89	88
$0,2$ $0,3$	0,46603	78	78
$0,3$ $0,4$	0,43667	64	58
$0,4$ $0,5$	0,40532	50	51
$0,5$ $0,6$	0,07607	36	36
$0,6$ $0,7$	0,05150	24	26
$0,7$ $0,8$	0,03265	15	14
$0,8$ $0,9$	0,01940	9	10
$0,9$ $1,0$	0,01447	5	7
au delà de $1,0$	0,01047	5	8

FIGURE 4 – Distribution des erreurs d'observation de la différence d'ascension droite entre le soleil et l'une des deux étoiles Atair et Procyon, données empiriques et données théoriques (Liagre, 1852)

d'ajustement de données à une courbe des erreurs (Pichard, 2005). Pour Quetelet, une persistance des causes se traduit par une tendance de la moyenne à être stable et inversement une modification de la moyenne indique une variation des causes, une évolution. Dans le cas de nombreuses petites causes indépendantes et de même influence, la résultante suit une loi des erreurs, comme par exemple pour la planche de Galton. Ceci sera confirmé par l'énoncé du théorème limite central. Les diverses expérimentations menées par Galton en anthropométrie, par le biologiste Weldon (1860-1906) sur les crevettes, crabes... et par d'autres ont permis de montrer que les distributions de fréquences des mesures effectuées dans les différents phénomènes ne suivent pas toujours une courbe des erreurs.

Pearson (1857-1936) cherche alors à caractériser la forme de différentes distributions empiriques. Il propose ensuite, dans les années 1893-1895, un système de distributions théoriques, c'est-à-dire un système de « fonctions qui sont des densités de probabilité et

## Émergence historique des lois à densité. Des pistes pour l'enseignement

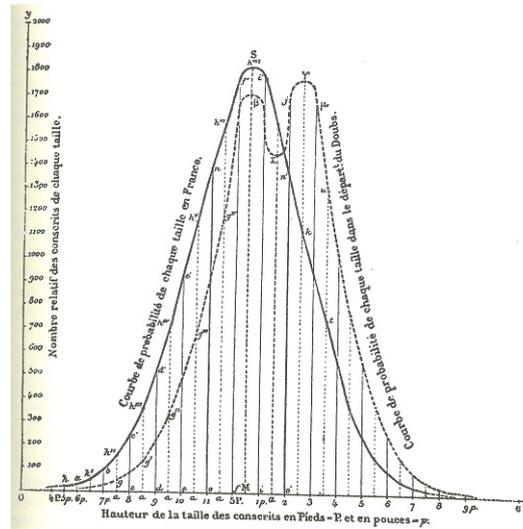


FIGURE 5 – Courbe des tailles des conscrits du département du Doubs en 1863 (Bertillon, 1876, cité dans Stigler, 1986, p. 217)

vérifient l'équation différentielle  $\frac{dy}{dx} = \frac{y(x+a)}{b_0 + b_1x + b_2x^2}$  sur un certain domaine  $D$  de valeurs de  $x$ . Les valeurs des paramètres  $a$ ,  $b_0$ ,  $b_1$ ,  $b_2$  sont déterminées par les moments d'ordre 1 à 4 de la loi ayant pour densité la fonction  $y$  » (Pichard, 2005, p. 242). Suivant les valeurs des paramètres  $a$ ,  $b_0$ ,  $b_1$ ,  $b_2$ , Pearson obtient différents types de distribution. En voici quelques exemples :

- Si  $b_0 \neq 0$ ,  $b_1 = b_2 = 0$ , les solutions de (\*) donnent les lois normales  $N(\mu, \sigma^2)$  avec  $b_0 = -\sigma^2$  et  $a = -\mu$  ;
- Si  $b_1 \neq 0$ ,  $b_2 = 0$ , les lois exponentielles et Gamma (Type III) ;
- Si  $b_0 \neq 0$ ,  $b_2 \neq 0$  tels que le trinôme  $b_0 + b_1x + b_2x^2$  est irréductible sur  $\mathbb{R}$ , les fonctions puissance de ce trinôme, dont les lois de Cauchy et de Student (Type IV) ;
- Si  $b_0 \neq 0$ ,  $b_2 \neq 0$  tels que le trinôme  $b_0 + b_1x + b_2x^2$  est réductible sur  $\mathbb{R}$ , les loi Bêta (Type I).

Ce système est enrichi par Pearson, puis complété par d'autres. Pearson fait donc apparaître de nouvelles lois d'un point de vue théorique. Nous pouvons aussi mentionner la loi du  $\chi^2$  (Type III) qui est utilisée dans le test statistique du  $\chi^2$  pour vérifier l'adéquation d'une distribution empirique à une loi de probabilité donnée. Toutes ces lois théoriques se retrouvent ensuite dans la pratique : la loi exponentielle fait son apparition pour les durées de vie, la loi de Pareto pour les distributions de revenus...

Dans ces travaux, nous voyons l'importance de la prise en compte de la distribution empirique pour vérifier la pertinence de la distribution théorique choisie. Effectivement, la loi des erreurs ou les autres lois ensuite sont là pour répondre à des problèmes concrets, d'où l'importance de s'assurer de la correspondance entre données empiriques et données

théoriques. Dans ces différents travaux apparaissent des courbes de fréquences (des diagrammes en bâtons reliés par une courbe) pour permettre ces vérifications. De plus, de nouvelles courbes émergent, non symétriques mais qui conservent les autres propriétés, notamment d'aire sous la courbe égale à 1. Ces nouvelles courbes permettent de résoudre d'autres problèmes que ceux liés aux erreurs de mesures.

### 2.3 La détermination empirique de lois : usage des histogrammes

Une fois les résultats théoriques vérifiés empiriquement, les résultats mathématiques sont diffusés à la communauté des astronomes et des géodésiens afin qu'ils puissent les appliquer à leurs données. Nous allons, dans cette partie, nous intéresser à la transposition didactique (Chevallard, 1985) proposée par Faye dans son *Cours d'astronomie et de géodésie* de 1881. Faye décide de présenter les choses différemment de celles exposées dans les cours de probabilités. Il explique son point de vue :

*Nous allons appliquer ici la méthode expérimentale, c'est-à-dire chercher par expérience, s'il existe une relation entre la probabilité d'une erreur et la grandeur de cette erreur, puis opérant par des mesures réellement faites, sur des écarts dûment constatés, nous tâcherons de découvrir graphiquement cette loi.*  
(Faye, 1928, p. 271)

Faye utilise le même exemple que Liagre (cf. figure 4). Sa démarche n'est cependant pas de comparer les valeurs empiriques et théoriques, mais de partir des valeurs observées pour déterminer empiriquement la loi que suivent les erreurs. Il s'agit d'un problème d'ajustement, qui reste un type de problème très important et d'actualité en mathématiques. Avant de commencer, en partant des données d'observations des erreurs sur des intervalles, Faye explique que la probabilité qu'une erreur soit comprise entre 0 et 0,1 n'est autre que la fréquence correspondante, donc le rapport

$$\frac{\text{nombre des erreurs de cet ordre}}{\text{nombre total des erreurs}}. \quad (2)$$

Il poursuit en disant que ces fractions ou probabilités suivent une certaine loi, qu'il faut déterminer graphiquement. Ensuite, il propose de représenter les données graphiquement sous forme d'un histogramme<sup>6</sup>. Cependant cette construction n'est pas directe et il explique en détail sa construction :

*Si l'on choisit un système d'axes des  $x$  et des  $y$ , et que l'on porte les erreurs sur l'axe des  $x$ , on ne saurait porter les probabilités correspondantes sur l'axe des  $y$ , car la probabilité d'une erreur  $x$  déterminée, c'est-à-dire comprise entre  $x$  et  $x + dx$ , doit être infiniment petite; elle n'est finie que si les limites de l'erreur diffèrent elles-mêmes d'une quantité finie. On est donc conduit à chercher graphiquement une courbe  $y = f(x)$  telle que les probabilités ci-dessus soient représentées par les aires comprises entre les couples d'ordonnées qui répondent aux abscisses 0 et 0,1, 0,1 et 0,2, ... Comme il y a autant d'erreurs*

<sup>6</sup>Le mot anglais « histogram » date de 1895 et sa traduction française normalisée, de 1948.

*Émergence historique des lois à densité. Des pistes pour l'enseignement*

*négatives que d'erreurs positives, cette courbe devra être symétrique de part et d'autre de l'axe des  $y$  ; en second lieu, entre les ordonnées relatives à  $x = 0$  et  $x = 0,1$ , l'aire devra être la même qu'entre les ordonnées de  $x = 0$  et  $x = -0,1$ , c'est-à-dire  $\frac{94}{2}$  ou 47. On obtient cette courbe en divisant l'axe des  $x$  en parties proportionnelles à 0, 1, 0, 2, 0, 3, ... et en construisant sur les intervalles de 0 à 0,1, de 0,1 à 0,2, ... des rectangles proportionnels aux nombres 47, 44, ... (Faye, 1928, p. 273)*

Faye met bien en évidence dans son discours que cette courbe (l'histogramme) ne porte pas la probabilité (en fait la fréquence) sur l'axe des ordonnées mais bien que la fréquence est proportionnelle à l'aire des rectangles. Cela permet ensuite de comprendre pourquoi c'est l'aire sous la courbe qui donne la probabilité et non l'image de la fonction. On voit apparaître une analogie forte entre fréquence et probabilité. Il poursuit la construction par le tracé de la courbe de la loi des erreurs en «*traçant un trait continu qui laisse, en dedans, de petites aires triangulaires équivalentes à celles qu'il laissera en dehors*» (cf. figure 6).

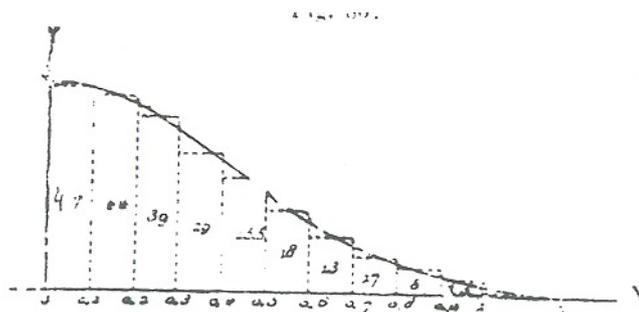


FIGURE 6 – *Moitié de la courbe représentant, par ses aires, la loi de probabilité des erreurs (Faye, 1928, p. 273)*

L'expression de la courbe est du type  $y = ae^{-h^2x^2}$ , où  $a$  et  $h$  sont des paramètres à déterminer. Faye justifie le fait de considérer une fonction strictement positive sur  $\mathbb{R}$ , bien qu'elle devrait être nulle en dehors des valeurs extrêmes, en expliquant qu'au-delà de ces abscisses, l'aire sous la courbe est négligeable donc que cela ne pose pas de problème. Il explique ensuite que la probabilité qu'une erreur soit comprise entre les limites finies  $x$  et  $x + \Delta x$ , c'est-à-dire le rapport

$$\frac{\text{aire de la courbe comprise entre les abscisses correspondantes}^7}{\text{aire totale de la courbe}} \quad (3)$$

<sup>7</sup>L'expression «aire de la courbe», reprise du texte de Faye, sous-entend l'aire sous la courbe. Par ailleurs, dans le texte de Faye, il est écrit au numérateur «aire de la courbe comprise entre les ordonnées», mais il s'agit bien des abscisses.

aura pour expression

$$\frac{\int_x^{x+\Delta x} ae^{-h^2x^2} dx}{\int_{-\infty}^{+\infty} ae^{-h^2x^2} dx}.^8 \quad (4)$$

La courbe choisie n'est effectivement pas normalisée, donc l'aire totale sous la courbe ne vaut pas nécessairement 1. Cependant l'intégrale du dénominateur étant connue, l'expression de la probabilité est ensuite donnée en fonction de  $h$  et  $x$ .

Faye donne ensuite un second exemple, non issu de la théorie des erreurs, tiré du *Traité des probabilités appliquées au tir* de Didion de 1858, où l'on retrouve à nouveau une courbe de la même famille.

Le cours de Faye est intéressant car, ne se voulant pas théorique, il propose une approche empirique où le point de départ sont les données statistiques, qui sont ensuite représentées sous forme d'histogramme. L'histogramme est un moyen de représenter les données qui, contrairement aux diagrammes en bâtons, permet de représenter la fréquence sur un intervalle, non pas grâce à l'axe des ordonnées mais bien grâce à l'aire des rectangles. Ensuite la courbe se déduit par un ajustement graphique.

### 3 Des pistes pour l'enseignement des lois à densité

L'étude historique et épistémologique présentée dans la partie précédente nous permet de dégager quelques points importants qui ont contribué à l'émergence de la notion de fonction de densité et plus généralement des lois à densité. Nous allons présenter les éléments qui ressortent de cette étude et nous dégagerons les pistes issues de ce travail pour introduire la notion de fonction de densité, pistes qui ont été exploitées lors d'expérimentations en classe de terminale S (Derouet, 2016). Bien entendu, ces seuls éléments ne sont pas suffisants et d'autres analyses (programme, manuels...) ont contribué à l'élaboration de la séquence d'enseignement testée mais, dans cet article, nous nous restreignons aux apports de l'étude historique et épistémologique.

#### 3.1 Des points clés à retenir de l'analyse historique

Les lois à densité émergent en réponse à des problèmes cherchant à décrire les erreurs de mesures : c'est donc bien en réponse à des problèmes concrets, dans le but d'interpréter des données statistiques sur des mesures. Cependant, dans un premier temps, l'approche est purement théorique. Les mathématiciens font des hypothèses en amont sur les distributions statistiques dont ils disposent, qui leur permettent de proposer des lois, dont la courbe de « facilité » vérifie certaines propriétés, notamment que l'aire sous la courbe permet de déterminer les probabilités. La courbe la plus rencontrée à l'époque est la courbe de Gauss qui ensuite sera jugée adéquate pour la loi des erreurs mais aussi pour modéliser de

<sup>8</sup>La double utilisation du  $x$  dans l'intégrale est celle trouvée dans le texte de Faye. Il faut cependant comprendre qu'il s'agit d'une abscisse donnée pour les bornes et d'une variable pour l'expression.

nombreux autres phénomènes. Ce choix de courbe est confirmé en confrontant la théorie avec les données empiriques.

Pour l'enseignement des probabilités à des astronomes et géodésiens, Faye expose dans son cours une approche différente. Dans cette transposition didactique, il propose une démarche inverse : le but est alors de partir des données, de passer par la représentation sous forme d'histogramme des fréquences associées aux données, pour ensuite trouver une courbe qui «épouse» bien l'histogramme. Dans ce cas, une place importante est laissée aux graphiques. Cette approche par les histogrammes fait apparaître le lien entre probabilité et calcul d'aire, et donc entre probabilités et calcul intégral.

Nous pouvons repérer le rôle clé de la statistique dans la découverte des lois à densité : dans un premier temps ce sont les données statistiques qui sont à l'origine de la recherche de modèles, et dans un second temps ces données permettent de valider les modèles ou encore de les déterminer empiriquement. Il existe une connexion forte entre statistique et probabilités. Le passage par les histogrammes permet de justifier le lien entre probabilité et aire sous la courbe de la fonction de densité, par analogie avec le lien entre fréquence et aire des rectangles de l'histogramme.

## **3.2 Des pistes didactiques pour l'introduction de la notion de fonction de densité dans l'enseignement**

Cette étude permet de mettre en évidence des éléments didactiques qui pourraient avoir leur intérêt dans l'introduction de la notion de fonction de densité dans les classes, en particulier en terminale S. Nous avons décidé de garder plusieurs remarques et pistes que nous détaillons ci-dessous.

### **3.2.1 Rôle de la fonction de densité dans l'introduction des lois à densité**

Historiquement, la fonction de densité est bien apparue avant la fonction de répartition. L'approche du programme de terminale S (MEN, 2011), qui mentionne la fonction de densité sans parler de fonction de répartition, est donc conforme à la construction historique des lois à densité. Bien que non entièrement rigoureuse pour une généralisation, une introduction des lois à densité avec la fonction de densité semble donc assez naturelle pour une première rencontre par les élèves.

### **3.2.2 Les lois à densité comme réponse à un problème**

La statistique a joué un rôle important dans l'émergence de la notion de fonction de densité. Les données statistiques de mesures observées, mais ensuite d'autres types de données (tailles de tour de poitrine, de tour de tête...), sont à l'origine de l'introduction des lois à densité. Ces lois sont donc arrivées en réponse à des problèmes concrets. Pour cette raison, il semble intéressant et important de garder à l'esprit le rôle des données statistiques comme point de départ dans le questionnement qui permet d'aboutir aux lois à densité. Une approche par des problèmes de modélisation probabiliste semble pertinente

et abordable en terminale S. Ce type d'approche permet effectivement de problématiser l'introduction de ces nouvelles notions.

Cependant, pour être plus proche des intérêts des élèves, il n'est peut-être pas souhaitable de partir de données sur les erreurs de mesures, mais de données portant sur d'autres phénomènes plus parlants pour les élèves. Nous pensons par exemple à l'étude de phénomènes naturels (éruptions volcaniques, crues...).

### 3.2.3 Une approche par ajustement

Une réflexion purement théorique n'est pas souhaitable en terminale. Elle est trop abstraite et s'appuie sur des notions statistiques complexes (maximum de vraisemblance...). En revanche, considérer un problème de modélisation en passant par un problème d'ajustement graphique est beaucoup plus envisageable. Le passage par les histogrammes, pour amener à la détermination de la fonction de densité, nous paraît pertinent et porteur de sens. Pour cela, il est cependant nécessaire de considérer des phénomènes pouvant être modélisés (tout du moins dans un premier temps) par des lois dont les fonctions de densité sont des fonctions connues et maîtrisées des élèves du niveau scolaire considéré. Cela leur permet d'être réellement acteur dans la recherche du modèle, par ajustement. Cette approche doit permettre de faire émerger les propriétés de la notion de fonction de densité. En effet, cela doit permettre de faire apparaître que la ressemblance entre le haut des rectangles de l'histogramme et la courbe ne suffit pas pour le choix de la fonction, mais que d'autres propriétés comme le fait que l'aire sous la courbe doit être égale à 1 sont indispensables.

Dans son cours, Faye (1928) ne s'adresse pas à des mathématiciens mais bien à des personnes désirant utiliser concrètement ses méthodes de détermination. Pour cette raison, cette approche semble plus problématisée et intéressante pour des élèves de lycée. Ce passage par le graphique avec l'histogramme nous semble intéressant, notamment pour faire émerger le rôle de l'aire sous la courbe. L'utilisation de l'histogramme nous semble donc une idée à retenir pour l'introduction de la notion de fonction de densité : il a l'avantage de donner du sens à la fonction de densité et notamment permet de construire la notion et ses propriétés caractéristiques, en liant notamment fréquence et aire des rectangles ce qui ensuite amène à relier probabilité et aire sous la courbe.

Le lien entre histogramme et courbe de densité est à mettre en correspondance avec le lien entre fréquence et probabilité qui est utilisé en classe dès le collège. Il s'agit implicitement de la loi faible des grands nombres, dont la formulation vulgarisée proposée (dans le cas fini) dans le document d'accompagnement du programme de 2001 de la classe de première S (MEN, 2001) est la suivante : «*Si on choisit  $n$  éléments d'un ensemble fini  $E$  selon une loi de probabilité  $P$ , indépendamment les uns des autres, alors la distribution des fréquences est proche de la loi de probabilité  $P$  lorsque  $n$  est grand*». En effet, si, par exemple, on simule plusieurs échantillons de taille suffisamment grande de variables aléatoires suivant une loi uniforme sur  $[0; 12]$ , on peut voir que le haut des rectangles des histogrammes est «proche» de la courbe représentative de la fonction  $x \mapsto 1/12$  qui n'est autre que la fonction de densité de la loi uniforme sur  $[0; 12]$  (cf. figure 7).

Cette approche ne nécessite pas des connaissances mathématiques plus importantes

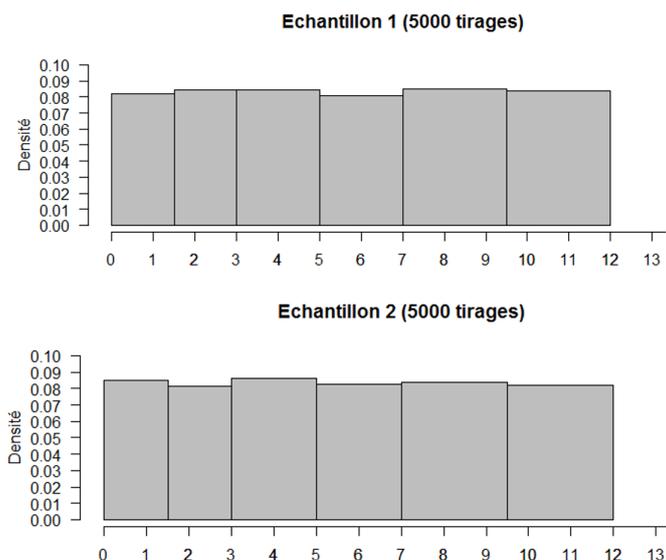


FIGURE 7 – *Histogrammes de deux échantillons de taille 5000 de variables aléatoires suivant une loi uniforme sur  $[0; 12]$ .*

que ce qui est attendu des élèves de terminale, mais en plus elle permet de valoriser l'objet mathématique « histogramme » qui est présent dans les programmes de mathématiques dès le collège, mais qui pourtant est très peu mis en avant (Derouet & Parzys, 2016 ; Roditi, 2009). Un autre avantage de cette approche est qu'elle nécessite l'utilisation des TICE. Le tableur ne permet pas la construction simple d'histogrammes de fréquences<sup>9</sup>, il n'est donc pas le bon outil technologique pour ce type de problème. En revanche, le logiciel *GeoGebra*, très utilisé en classe, est très simple de prise en main pour la construction d'histogrammes et ensuite pour la recherche de courbe de densité par ajustement. D'autres logiciels, comme R, peuvent aussi être utilisés mais sont beaucoup moins « transparents » pour les élèves.

### 3.3 Des pistes qui se retrouvent dans d'autres recherches

Ce point de vue historique et épistémologique renforce des éléments déjà mis en avant dans d'autres recherches en didactique des mathématiques. Nous citerons les travaux de Batanero *et al.* (2004) dans lesquels ils ont mis en évidence l'importance de travailler sur les représentations graphiques pour aider les étudiants à la compréhension de la loi normale. Les auteurs insistent : « *il est important que les étudiants comprennent les concepts basiques comme les notions de probabilité, de courbe de densité [...] et d'histogramme avant de commencer l'étude de la loi normale*<sup>10</sup> » (p. 275). Les principales difficultés qu'ils ont repérées chez les étudiants sont des difficultés à percevoir l'intérêt d'un modèle pour décrire des données empiriques, des difficultés à interpréter des aires d'histogrammes de fréquences, des difficultés à interpréter les probabilités sous la courbe... Ces difficultés sont

<sup>9</sup>Le tableur appelle « histogramme » ce qui correspond en fait à un diagramme en bâtons.

<sup>10</sup>Traduit par nos soins.

effectivement prises en compte et travaillées dans les pistes soulevées.

## 4 En guise de conclusion : vers une ingénierie didactique

Dans cet article, nous avons retracé des éléments de l'émergence historique des lois à densité, à travers la théorie des erreurs, qui ont permis de mettre en évidence des pistes à exploiter pour l'enseignement de ces notions. Nous rappelons quelques points qui nous semblent importants :

- les lois à densité doivent être introduites comme réponse à un problème ;
- les données statistiques de départ doivent être parlantes pour les élèves ;
- l'appui sur les représentations graphiques est important : la représentation sous forme d'histogramme permet, par ajustement, de faire émerger la courbe de la fonction de densité et ses propriétés.

Suite à cette analyse historique et épistémologique, enrichies d'autres analyses, les pistes dégagées nous ont ensuite conduit à la conception d'une ingénierie didactique (Artigue, 1988) dont l'objectif est de permettre aux élèves de construire la notion de fonction de densité (et ses propriétés), en articulant à la fois l'introduction des lois à densité et du calcul intégral, en terminale S (Derouet, 2016) pour donner du sens à l'égalité

$$P(a \leq X \leq b) = \int_a^b f(x) dx.$$

En guise de conclusion, nous donnons quelques éléments sur un des problèmes conçus. Il s'agit d'un problème d'estimation du temps d'attente entre deux éruptions du volcan Aso (Japon), à partir des données des années des 139 éruptions du volcan Aso entre le XIII<sup>e</sup> et le XXI<sup>e</sup> siècle<sup>11</sup>. Ces données peuvent être représentées sous forme d'histogramme (cf. figure 8). Un travail d'ajustement peut ensuite être effectué, en prenant en compte la ressemblance entre l'histogramme et la courbe, mais aussi les propriétés de la fonction de densité (positivité, aire sous la courbe égale à 1). Dans ce cas, une courbe de type exponentielle décroissante peut être proposée. Bien entendu, pour la validation du modèle, des considérations spécifiques aux éruptions volcaniques seraient ensuite à prendre en compte, cependant ce type de travail n'est pas attendu des élèves.

Ce problème de modélisation reprend la démarche proposée par Faye en 1928 tout en s'adaptant aux contraintes de l'enseignement secondaire en mettant en jeu des fonctions connues et maîtrisées des élèves de terminale S, mais aussi en étudiant des phénomènes plus attractifs que les erreurs de mesures.

Un autre des enjeux de la classe de terminale S est d'introduire le calcul d'aire sous une courbe à l'aide de l'intégrale. Cette notion d'aire sous la courbe (et donc d'intégrale) se retrouve aussi dans l'introduction des lois à densité, où s'ajoute de surcroît le cas des

<sup>11</sup>Les données se trouvent à l'adresse suivante : <http://www.volcano.si.edu/volcano>.

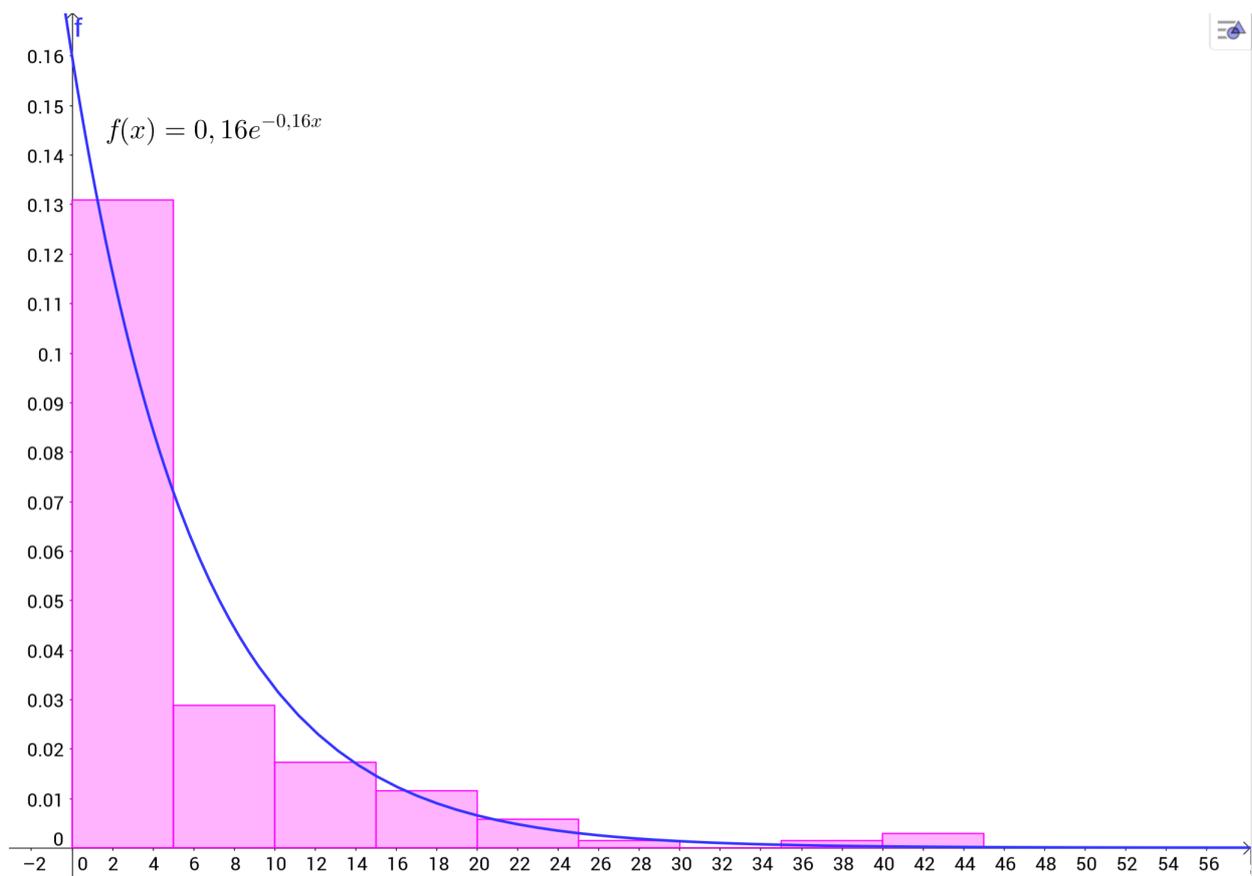


FIGURE 8 – *Histogramme pour des classes d'amplitude 5 ans des temps d'attente entre deux éruptions du volcan Aso depuis le XIII<sup>e</sup> siècle.*

bornes infinies. Nous avons testé l'hypothèse que l'introduction des lois à densité prend d'autant plus de sens quand elle est conjointe à celle du calcul intégral (Derouet, 2016).

Nous pourrions aussi envisager, une fois la notion de fonction de densité introduite, de faire rencontrer aux élèves des fonctions de densité moins «classiques». En figure 9 se trouve l'histogramme des temps d'attente (en minutes) entre deux jets d'eau du geyser Old Faithful<sup>12</sup>, aux États-Unis. Dans ce cas, c'est une distribution bimodale (comme pour le cas de la taille des conscrits) qui apparaît.

Ces propositions faites dans le cadre de l'enseignement secondaire pourraient tout à fait être transférées dans l'enseignement supérieur. Il serait notamment intéressant de comparer les approches des étudiants suivant qu'ils rencontrent ou non pour la première fois ces notions.

<sup>12</sup>Les données se trouvent à l'adresse suivante : <http://www.stat.cmu.edu/~larry/all-of-statistics/=data/faithful.dat>.

C. Derouet

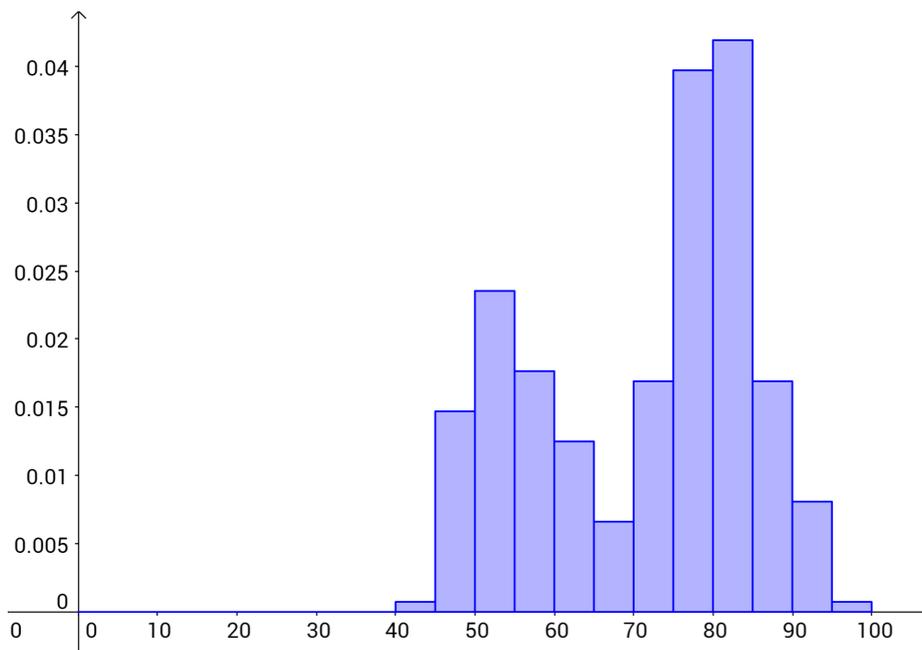


FIGURE 9 – *Histogramme pour des classes d'amplitude 5 minutes des temps d'attente (en minutes) entre deux jets d'eau du geyser Old Faithful.*

## Références

- [1] Armatte, M. (2004), La théorie des erreurs (1750-1820) : enjeux, problématiques, résultats. In E. Barbin & J.-P. Lamarche (Eds.), *Histoires de probabilités et de statistiques* (pp. 141-160), Ellipses Édition, Paris.
- [2] Artigue, M. (1988), Ingénierie didactique, *Recherches en Didactique des Mathématiques*, **9**(3), 281-308.
- [3] Batanero, C., L. M. Tauber & V. Sánchez (2004), Students' reasoning about the normal distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 257-276), Springer, Dordrecht.
- [4] Bru, B. (2006), La courbe de Gauss ou le théorème de Bernoulli raconté aux enfants, *Mathématiques & sciences humaines – Mathematics & social sciences*, **175**, 5-23.
- [5] Chevallard, Y. (1985), *La transposition didactique : du savoir savant au savoir enseigné*, La Pensée sauvage, Grenoble.
- [6] Derouet, C. (2016), *La fonction de densité au carrefour entre probabilités et analyse. Étude de la conception et de la mise en oeuvre de tâches d'introduction articulant lois à densité et calcul intégral*, Thèse de doctorat, Université Paris Diderot.
- [7] Derouet, C. & B. Parzysz (2016), *How can histograms be useful for introducing continuous probability distributions ?*, *ZDM – Mathematics Education*, **48**(6), 757-774.

- [8] Dorier, J. (2000), Recherches en histoire et en didactique des mathématiques sur l'algèbre linéaire. Perspectives théoriques sur leurs interactions, *Les cahiers du laboratoire Leibniz*, **12**.
- [9] Droysbeke, J.-J. & P. Tassi (1990), *Histoire de la statistique*, Presses Universitaires de France, Paris.
- [10] Lefort, X. & A. Boyé (1996), De Cassini à Gauss : du calcul d'erreurs aux probabilités. In *Contribution à une approche historique de l'enseignement des mathématiques* (pp. 239-258), IREM de Franche-Comté, Besançon.
- [11] Ministère de l'Éducation nationale (MEN) (2011), Programme de l'enseignement spécifique et de spécialité de mathématiques. Classe terminale de la série scientifique, *Le B.O.*
- [12] Ministère de l'Éducation nationale (MEN) & Direction de l'Enseignement Scolaire (2001), *Mathématiques. Classe de première des séries générales*, Centre national de documentation pédagogique, Paris.
- [13] Parzysz, B. (2013), La longue genèse de la loi normale, *Le Bulletin Vert*, **502**, 29-40.
- [14] Pichard, J.-F. (2005), Théorie des erreurs, courbes en cloche et normalité. In B. Chaput & M. Henry (Eds.), *Statistique au lycée. Volume 1 : Les outils de la statistique* (pp. 219-246), Paris.
- [15] Robinet, J. (1984), *Ingénierie didactique de l'élémentaire au supérieur*, Thèse de doctorat, Université Paris VII.
- [16] Roditi, E. (2009), L'histogramme : à la recherche du savoir à enseigner, *Spirale-Revue de recherches en éducation*, **43**, 129-138.
- [17] Stigler, S. M. (1986), *The history of statistics : The measurement of uncertainty before 1900*, Harvard University Press.

## Annexe 1

Range (seconds)	Frequency of errors		
	Observed	Calculated	Recalculated
A. 300 Observations of declinations			
0.0–0.4	66	65	64.63
0.4–0.8	58	60	60.01
0.8–1.2	55	53	51.72
1.2–1.6	28	41	41.39
1.6–2.0	27	30	30.75
2.0–2.4	23	21	21.21
2.4–2.8	10	13	13.58
2.8–3.2	15	8	8.08
3.2–3.6	8	5	4.46
3.6–4.0	4	2	2.28
Above 4.0	6	2	—
B. 300 Observations of right ascensions			
0.0–0.1	114	107	107.22
0.1–0.2	84	87	86.80
0.2–0.3	53	57	56.87
0.3–0.4	24	30	30.14
0.4–0.5	14	13	12.91
0.5–0.6	6	5	4.47
0.6–0.7	3	1	1.25
0.7–0.8	1	0	0.28
0.8–0.9	1	0	0.05
C. 470 Observations of right ascensions			
0.0–0.1	94	95	94.39
0.1–0.2	88	88	88.50
0.2–0.3	78	78	77.81
0.3–0.4	58	64	64.15
0.4–0.5	51	49	49.58
0.5–0.6	36	35	35.94
0.6–0.7	26	24	24.42
0.7–0.8	14	16	15.56
0.8–0.9	10	9	9.30
0.9–1.0	7	5	5.21
Above 1.0	8	5	—

## Annexe 2

Table n° 3.

VALEURS DE L'INTÉGRALE DÉFINIE  $P_2 = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$ , POUR DES  
VALEURS DE  $t$  EXPRIMÉES EN FONCTION DE  $\rho$  PRIS POUR UNITÉ.

$\frac{t}{\rho}$	$\frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$	Différences	$\frac{t}{\rho}$	$\frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$	Différences
0,0	0,000	54	2,5	0,908	13
0,1	0,054	53	2,6	0,921	10
0,2	0,107	53	2,7	0,934	10
0,3	0,160	53	2,8	0,944	9
0,4	0,213	51	2,9	0,950	7
0,5	0,264	50	3,0	0,957	6
0,6	0,314	49	3,1	0,963	6
0,7	0,363	48	3,2	0,969	5
0,8	0,411	45	3,3	0,974	4
0,9	0,456	44	3,4	0,978	4
1,0	0,500	42	3,5	0,982	3
1,1	0,542	40	3,6	0,985	2
1,2	0,582	37	3,7	0,987	3
1,3	0,619	36	3,8	0,990	1
1,4	0,655	33	3,9	0,991	2
1,5	0,688	31	4,0	0,993	1
1,6	0,719	29	4,1	0,994	1
1,7	0,748	27	4,2	0,995	1
1,8	0,775	25	4,3	0,996	1
1,9	0,800	23	4,4	0,997	1
2,0	0,823	20	4,5	0,998	0
2,1	0,843	19	4,6	0,998	0
2,2	0,862	17	4,7	0,998	1
2,3	0,879	16	4,8	0,999	0
2,4	0,895	13	4,9	0,999	0
2,5	0,908		5,0	0,999	0

Cette table est indépendante de la précision des observations : elle donne la probabilité que l'erreur, pour une espèce quelconque d'observations, ne dépasse pas une certaine valeur exprimée en fonction de l'erreur probable.

Elle montre que, sur 1000 erreurs, il en reste 54 au-dessous de 0,1 de l'erreur probable ; 107 au-dessous de 0,2, etc. En d'autres termes, on peut parier 54 contre 946 que l'erreur que l'on commettra, dans une espèce quelconque d'observations, sera moindre que 0,1 de l'erreur probable ; 107 contre 893 qu'elle sera moindre que 0,2 de l'erreur probable, etc.