

Le bulletin de l'APMEP - N° 552

# AU FIL DES MATHS

de la maternelle à l'université

Avril, mai, juin 2024

**Automat(h)ismes**



# APMEP

Association des Professeurs de Mathématiques de l'Enseignement Public

# ASSOCIATION DES PROFESSEURS DE MATHÉMATIQUES DE L'ENSEIGNEMENT PUBLIC

26 rue Duméril, 75013 Paris

Tél. : 01 43 31 34 05

Courriel : secretariat-apmep@orange.fr - Site : <https://www.apmep.fr>

Présidente d'honneur : Christiane ZEHREN

*Au fil des maths*, c'est aussi une revue numérique augmentée :

<https://afdm.apmep.fr>



Les articles sont en accès libre, sauf ceux des deux dernières années qui sont réservés aux adhérents *via* une connexion à leur compte APMEP.

Si vous désirez rejoindre l'équipe d'*Au fil des maths* ou bien proposer un article, écrivez à [aufildesmaths@apmep.fr](mailto:aufildesmaths@apmep.fr)

Annonces : pour toute demande de publicité, contactez Mireille GÉNIN [mcgenin@wanadoo.fr](mailto:mcgenin@wanadoo.fr)

À ce numéro est joint le BGV n° 236  
spécial « Journées Nationales »

## ÉQUIPE DE RÉDACTION

**Directrice de publication** : Claire PIOLTI-LAMORTHE.

**Responsable coordinatrice de l'équipe** : Cécile KERBOUL.

**Rédacteurs** : Vincent BECK, François BOUCHER, Richard CABASSUT, Séverine CHASSAGNE-LAMBERT, Frédéric DE LIGT, Mireille GÉNIN, Cécile KERBOUL, Valérie LAROSE, Alexane LUCAS, Lise MALRIEU, Marie-Line MOUREAU, Serge PETIT, Thomas VILLEMONTÉIX, Christine ZELTY.

« **Fils rouges** » numériques : Gwenaëlle CLÉMENT, François COUTURIER, Jonathan DELHOMME, Nada DRAGOVIC, Fanny DUHAMEL, Laure ÉTEVEZ, Marianne FABRE, Yann JEANRENAUD, Armand LACHAND, Lionel PRONOST, Agnès VEYRON.

**Illustrateurs** : Stéphane FAVRE-BULLE, Pol LE GALL, Olivier LONGUET.

**Équipe T<sub>E</sub>Xnique** : Sylvain BEAUVOIR, Laure BIENAIMÉ, Isabelle FLAVIER, Philippe PAUL, François PÉTIARD, Guillaume SEGUIN, Sébastien SOUCAZE, Sophie SUCHARD.

**Maquette** : Olivier REBOUX.

**Correspondant Publimath** : François PÉTIARD.

**Votre adhésion à l'APMEP vous abonne automatiquement à *Au fil des maths*.**

Pour les établissements, le prix de l'abonnement est de 60 € par an.

La revue peut être achetée au numéro au prix de 15 € sur la boutique en ligne de l'APMEP.

Mise en page : François PÉTIARD

Dépôt légal : juin 2024. ISSN : 2608-9297.

Impression : iLLiCO by L'ARTÉSIENNE

ZI de l'Alouette, Rue François Jacob, 62800 Liévin



# La loi de Benford

*La loi de Benford régit la répartition des premiers chiffres de séries statistiques invariantes par changement d'échelles et fournit des résultats assez contre-intuitifs sur cette répartition. Jean Lefort nous en propose une exploration à la main.*

Jean Lefort

## Un curieux paradoxe

Imaginons que nous relevions plusieurs centaines de données numériques entières. Certaines commencent par le chiffre 1, d'autres par 2, etc. Aucune ne commence par 0. Si, par exemple, il s'agit de la taille en centimètres d'individus adultes, il est certain que le chiffre 1 commencera la quasi-totalité des mesures. Dans d'autres cas, on peut s'attendre à une répartition quasi égale entre les neuf chiffres. Or ce n'est pas ce que l'on constate dans les cas suivants :

1. les prix de tous les articles (exprimés pour simplifier en centimes) dans un supermarché ;
2. la longueur, exprimée en kilomètres, de tous les fleuves et toutes les rivières de France ;
3. la population de toutes les communes de France.

*Dans tous ces cas on trouve près de deux fois plus de données qui commencent par 1 que par 2 puis une diminution de plus en plus lente pour arriver à environ 5 % qui commencent par 9. Bizarre !*

Cet article se penche sur les calculs qui peuvent être réalisés pour comprendre le phénomène de décroissance ainsi que les pourcentages obtenus pour chacun des chiffres. Pour comprendre à quel type de données il s'applique et le sens de l'invariance par changement d'unité, on pourra consulter l'article [1] et le chapitre 1 du *Théorème du parapluie* [2].

## Un exemple

On trouve facilement sur l'internet la population des communes d'un département donné ou d'une région donnée. Il est alors facile de trier les données en fonction du premier chiffre.

En bas de cette page (table 1) se trouve le résultat concernant le recensement de 2015 pour les 883 communes alsaciennes. On remarque que, à l'exception des nombres commençant par 8 ou 9, il y a une décroissance très nette. Il est vraisemblable que l'exception est due à la faiblesse des effectifs concernés.

### La décroissance

Dans ce qui suit nous nous intéressons plus particulièrement à la répartition des prix dans une grande surface et nous ferons donc l'hypothèse que cette répartition est indépendante de l'unité monétaire choisie. Nous sommes conscients que bien des prix sont baissés de quelques centimes, cinq ou dix, pour afficher 9,95 € au lieu de 10 € ou 14,90 € au lieu de 15 €. Difficile de savoir comment cela joue et avec quelle importance sur la répartition du chiffre le plus à gauche, d'autant plus, et nous l'avons vu lors du passage à l'euro, que l'entreprise adopte la même stratégie avec la nouvelle unité.

Chiffre	1	2	3	4	5	6	7	8	9
Nombre	226	131	119	102	79	77	61	43	45
%	25,59	14,84	13,48	11,55	8,95	8,72	6,91	4,87	5,10

Table 1.



Imaginons que l'on multiplie ou divise tous les prix par un même nombre. Que va-t-il se passer pour les prix ? Pour simplifier le calcul, supposons que tous les prix soient multipliés par 2 (on les exprime dans une monnaie qui vaut deux fois moins).

Ceux qui commençaient par 1 vont commencer par 2 ou 3.  
 Ceux qui commençaient par 2 vont commencer par 4 ou 5.  
 Ceux qui commençaient par 3 vont commencer par 6 ou 7.  
 Ceux qui commençaient par 4 vont commencer par 8 ou 9.

Notons  $p(x)$  la probabilité qu'un prix commence par  $x$ . On peut également s'intéresser au cas où  $x$  s'exprime avec deux chiffres ou plus (en cherchant la probabilité que l'écriture d'un prix commence par 23, par exemple). Comme nous l'avons annoncé, nous faisons l'hypothèse que la répartition des prix est indépendante de l'unité monétaire choisie. Le raisonnement précédent prouve que :

$$\begin{aligned} p(1) &= p(2) + p(3) ; p(2) = p(4) + p(5) ; \\ p(3) &= p(6) + p(7) ; p(4) = p(8) + p(9) ; \quad (1) \\ p(1) &= p(5) + p(6) + p(7) + p(8) + p(9) \end{aligned}$$

La dernière égalité résultant des quatre autres. On en déduit que  $p(1) > p(2)$ .

Effectuons le même raisonnement en multipliant tous les prix par 1,5 :

Ceux qui commençaient par 1 vont commencer par 15, 16, ..., 28, 29.  
 Ceux qui commençaient par 2 vont commencer par 3, 40, 41, 42, 43 ou 44.  
 Etc. On obtient alors la relation :

$$p(2) = p(3) + p(40) + p(41) + p(42) + p(43) + p(44).$$

On en déduit que  $p(2) > p(3)$ .

Des raisonnements analogues avec des multiplications par  $\frac{4}{3}$ ,  $\frac{5}{4}$ ,  $\frac{6}{5}$ ,  $\frac{7}{6}$ ,  $\frac{8}{7}$ , et  $\frac{9}{8}$  permettent de conclure que :

$$p(1) > p(2) > p(3) > p(4) > p(5) > p(6) > p(7) > p(8) > p(9).$$

On pourrait prolonger les inégalités pour 10, 11, 12, etc.

Le raisonnement que nous venons de faire à propos de la série statistique des prix d'un supermarché, nous le faisons de la même manière sur la série statistique des longueurs des cours d'eau, longueur que nous pouvons mesurer en kilomètres, en miles nautiques ou en toute autre unité qui nous plaira.

### Passage de $x$ à $2x$

Nous avons vu que dans le cas où  $x$  vaut 1, 2, 3 ou 4, nous avons  $p(x) = p(2x) + p(2x + 1)$ . En prenant pour  $x$  des valeurs à 1 chiffre ou 2 chiffres ou plus, cette formule se généralise. Prenons un exemple.

Considérons  $x = 23$  et doublons tous les prix. Pour  $230 \leq x < 235$  on a  $460 \leq 2x < 470$ . Puis pour  $235 \leq x < 240$  on a  $470 \leq 2x < 480$ , ce qui prouve qu'on a  $p(23) = p(46) + p(47)$ .

On n'oublie pas qu'il s'agit de probabilités et par suite, on a la relation  $\sum_{n=1}^9 p(n) = 1$ .

Il semble bien qu'à partir de ces deux relations les valeurs de  $p$  soient déterminées comme on peut s'y atteler par approximations successives. Dans un premier temps, considérons une fonction  $g$  décroissante qui satisfait aux deux conditions :  $g(1) = 100$  et  $g(x) = g(2x) + g(2x + 1)$ . On trouve à peu près :

$$\begin{aligned} g(1) &= 100 ; & g(2) &= 60 ; & g(3) &= 40 ; \\ g(4) &= 35 ; & g(5) &= 25 ; & g(6) &= 21 ; \\ g(7) &= 19 ; & g(8) &= 18 ; & g(9) &= 17 \end{aligned}$$

qu'il faut ramener en pourcentage en divisant par  $335 = 100 + 60 + 40 + 34 + 25 + 21 + 19 + 18 + 17$  ce qui permet alors d'avoir des probabilités.

On peut proposer d'autres répartitions satisfaisant aux équations (1), mais on se rend vite compte, surtout si on prolonge la série pour les deux premiers chiffres à gauche qu'on est de plus en plus contraint si l'on veut respecter la décroissance.





### La loi de Benford

Voyons comment nous pouvons être plus précis en faisant une hypothèse supplémentaire. Supposons, dans un premier temps, qu'il existe une fonction  $f$  telle que  $p(x) = f(x + 1) - f(x)$ . En utilisant cette formule, il vient :

$$\begin{aligned} p(x) &= f(x+1) - f(x) = p(2x) + p(2x+1) \\ &= (f(2x+1) - f(2x)) + (f(2x+2) - f(2x+1)) \\ &= f(2x+2) - f(2x). \end{aligned}$$

D'où  $f(x) - f(2x) = f(x + 1) - f(2x + 2)$ . Ceci prouve que la variation de la fonction  $f$  entre un nombre et son double est indépendante de la valeur de ce nombre. En fait cela n'est valable que pour les entiers. Une famille de fonctions vient immédiatement à l'esprit, les logarithmes. Ce ne sont pas les seules mais ce sont les plus simples. On peut, par exemple, ajouter à un logarithme une fonction périodique  $\omega$  de période 1, puisque nous ne travaillons que sur  $\mathbb{N}^*$ . Un calcul élémentaire permet de vérifier qu'une telle fonction  $f(x) = a \ln(x) + \omega(x)$  satisfait à  $f(x) - f(2x) = f(x + 1) - f(2x + 2)$ .

Alors :  $f(x) = a \ln(x) + \omega(x)$  et

$$p(x) = a \ln(x+1) - a \ln(x) = a \ln\left(\frac{x+1}{x}\right) = a \ln\left(1 + \frac{1}{x}\right).$$

Il faut maintenant nous souvenir que  $p$  représente une probabilité et que  $\sum_{n=1}^9 p(n) = 1$ . Donc :

$$(a \ln(2) - a \ln(1)) + (a \ln(3) - a \ln(2)) + \dots + (a \ln(10) - a \ln(9)) = a \ln(10) = 1.$$

Ceci détermine  $a$  et prouve que  $a \ln(x) = \log(x)$  (logarithme décimal) et  $p(x) = \log\left(1 + \frac{1}{x}\right)$ , d'où :

$$\begin{aligned} p(1) &= 0,301 ; & p(2) &= 0,176 ; & p(3) &= 0,125 ; \\ p(4) &= 0,097 ; & p(5) &= 0,079 ; & p(6) &= 0,067 ; \\ p(7) &= 0,058 ; & p(8) &= 0,051 ; & p(9) &= 0,046. \end{aligned}$$

On pourra comparer avec le premier exemple sur la population des communes alsaciennes ou sur la tentative d'approximation que nous avons faite un peu plus haut.

### Remarque

Un mathématicien peut s'intéresser à l'écriture des nombres dans une autre base que dix, soit  $b$ . Alors la fonction  $p$  qui intervient est définie par  $p(x) = \log_b\left(1 + \frac{1}{x}\right)$ .

### Conclusion

Comme nous l'avons annoncé au début, le raisonnement qui vient d'être fait à propos des prix s'applique également aux longueurs des cours d'eau que l'on peut mesurer dans différentes unités. C'est plus difficile pour des populations sauf à les regrouper par paquets de  $n$  unités. Cette loi s'applique également pour le premier chiffre des puissances successives de 2 [3].

La loi de Benford qui annonce que la répartition du ou des premiers chiffres  $x$  d'une série de données suffisamment importante suit une loi de probabilité définie par  $p(x) = \log\left(1 + \frac{1}{x}\right)$  s'applique à toutes les données dont on peut raisonnablement penser que leur répartition est indépendante du choix de l'unité utilisée pour les mesurer. Ainsi cette loi a été utilisée pour vérifier la compatibilité de grandes entreprises ou d'États avec mise en évidence de fraude dans certains cas.

### Références

- [1] Françoise Valette-Duchêne. *Loi de Benford, des mathématiques élémentaires pour débusquer des fraudes ou des erreurs en économie (ou ailleurs)*. 2012.
- [2] Mickaël Launay. *Le théorème du parapluie, chapitre 1*. Flammarion, 2019.
- [3] V. Arnold et A. Avez. *Ergodic problems of classical mechanics*. Benjamin, 1968.



Jean Lefort, retraité, ancien animateur de l'IREM de Strasbourg, ayant enseigné en lycée puis en prépa technologique, a beaucoup œuvré pour la vulgarisation des maths.

Il est membre de l'APMEP depuis 1968.

[jlefort.apmep@wanadoo.fr](mailto:jlefort.apmep@wanadoo.fr)

© APMEP Juin 2024







# Sommaire du n° 552



## Automat(h)ismes

### Éditorial

1

Fabrication de très grandes boîtes... la suite !

*Florence Soriano-Gafiuk & Manuella Freyermuth* ..... 59

### Opinions

✦ La parole au groupe « Fondamentaux et Automatismes »

*Groupe « Fondamentaux et Automatismes »* ..... 3

Croisements de points de vue sur la mesure

*Aurélie Chesnais & Valérie Munier* ..... 8

✦ Automatismes ou automathismes ?

*Éric Trouillot* ..... 21

✦ Des Mises En TRAIN pour bien démarrer

*Claire Piolti-Lamorthe & Sophie Roubin* ..... 26

### Avec les élèves

✦ Des rituels en collège

*Lydie El-Halougi* ..... 35

Double vue

*Jean-Christophe Deledicq* ..... 39

✦ MathsMentales

*Sébastien Coge* ..... 41

✦ MathALÉA : du nouveau !

*Ève Chambon, Lydie El Halougi & Stéphane Guyon*... 45

✦ Automatismes : un peu, beaucoup, passionnément...

*Céline Bruel & Élise Locatelli* ..... 50

### Ouvertures

La loi de Benford

*Jean Lefort* ..... 56

La Grande Aventure des maths

*C. Sakarovitch, G. Mulsant & M. Andler* ..... 65

Des bulles aux polyèdres

*Richard Cabassut* ..... 71

### Récréations

Au fil des problèmes

*Frédéric de Ligt* ..... 75

Des problèmes dans nos classes

*Valérie Larose* ..... 77

### Au fil du temps

Hommage à Guy Brousseau

*Éric Barbazo* ..... 79

Le CDI de Marie-Ange

*Marie-Ange Ballereau* ..... 81

Matériaux pour une documentation ..... 83

Les fichiers *Evariste* : toujours d'actualité !

*Jean Fromentin & Nicole Toussaint* ..... 87

Des étudiants aux Journées Nationales à Rennes

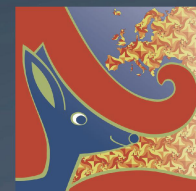
*Christophe Rivière* ..... 90

Mes premières Journées Nationales

*Mathieu Boutier* ..... 94



CultureMATH



# APMEP

www.apmep.fr