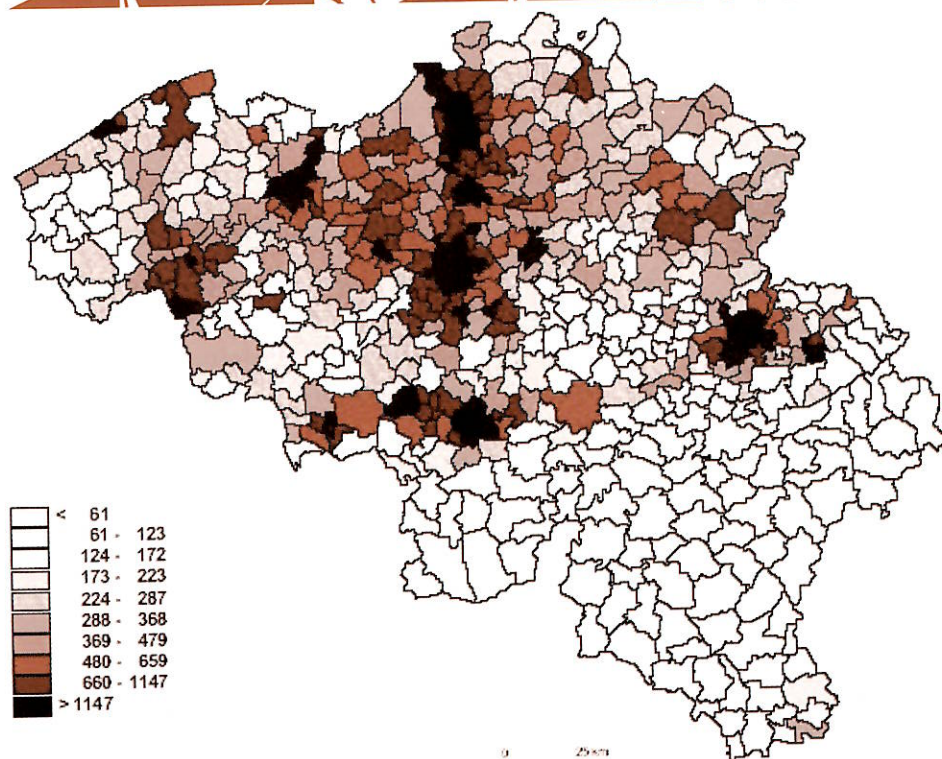


Novembre 2005 – 27<sup>e</sup> année – N 112 S



# Math-Jeunes

La Statistique





## Actualités

### Apprenti Géomètre

Vous connaissez Cabri-Géomètre ? **Apprenti Géomètre** est un logiciel de la même famille ayant des spécificités uniques, notamment des formes géométriques préprogrammées et la possibilité de découper et fusionner des figures. Téléchargeable gratuitement sur [www.enseignement.be/crem](http://www.enseignement.be/crem)

**Les 31<sup>e</sup> Olympiades Mathématiques Belges**

En 2006, les Olympiades se dérouleront selon le calendrier suivant :

- Éliminatoires : Mercredi 18 janvier 2006.
- Demi-finales : Mercredi 8 mars 2006.
- Finales : Mercredi 26 avril 2006.

- Proclamation : Samedi 13 mai 2006.

L'inscription s'effectue par l'intermédiaire de l'école.

### Le palimpseste d'Archimède

Un *palimpseste* est un manuscrit qui a été utilisé deux fois : après qu'un premier texte y ait été écrit, le support — parfois plusieurs siècles plus tard — est lavé et réemployé. C'est ainsi que le seul manuscrit qui nous soit parvenu du traité d'Archimède sur « La méthode des théorèmes mécaniques », est un palimpseste datant du <sup>x<sup>e</sup></sup> siècle, réutilisé au <sup>xii<sup>e</sup></sup> par un moine en vue d'y transcrire un livre de prières.

Identifié en 1907, environ 80% du texte d'Archimède a pu jusqu'à présent être reconstitué. Inutile de dire que les mathématiciens et les historiens des mathématiques souhaitent pouvoir combler les lacunes ! Uwe Bergmann, un physicien de l'Université de Stanford a constaté qu'il pouvait utiliser un accélérateur de particules pour détecter d'infimes parcelles de fer provenant de l'encre utilisée par le scribe. Il vient ainsi de déchiffrer trois des pages manquantes et espère terminer le travail en trois ou quatre années.

# Sommaire

S. Trompler, Edmund Halley	4
C. Festraets, Olympiades	5
G. Noël, Elle tremble, elle tremble !	6
P. Tilleuil, Tout droit dans le nuage !	10
J.-M. Desbonnez, Calcul statistique d'une aire	16
D. Justens, Histoire et mathématiques des tables de mortalité	20
Y. Noël-Roch, Jeux	28
N. Miewis, Rallye problèmes	31

En couverture : un diagramme statistique représentant, commune par commune, la densité de la population en Belgique. Plus la densité est importante, plus la teinte est foncée.



Un statisticien belge de renommée internationale.

**Lambert Adolphe QUÉTELET (1796-1874)**

Adolphe Quételet fut diplômé docteur à l'université de Gand en 1819 pour une thèse sur les sections coniques. Il enseigna ensuite à Bruxelles et, en 1823, partit étudier l'astronomie à l'Observatoire de Paris. Il y eut les plus grands maîtres tels Arago. Il étudia les probabilités sous la direction de Joseph Fourier et Pierre Laplace.

Revenu à Bruxelles en 1833 à la demande du gouvernement belge, Quételet étudia les données statistiques, géophysiques et météorologiques et établit des méthodes pour la comparaison et l'évaluation de ces données. En 1835 il publia « *Sur l'homme et le développement de ses facultés, essai d'une physique sociale* ». Il y présentait sa conception de l'« homme moyen ».

À partir de la loi normale, il fit des travaux très controversés par les sociologues de son époque sur le lien entre les criminels et leur statut social. En 1853, il organisa la première conférence internationale de statistique. On lui doit également la définition de la mesure de l'obésité encore utilisée de nos jours : l'indice de Quételet (IQ) d'un individu est le quotient du poids (en kilogrammes) par le carré de la taille (en mètres). Si IQ est supérieur à 30, la personne est obèse.

S. Trompler



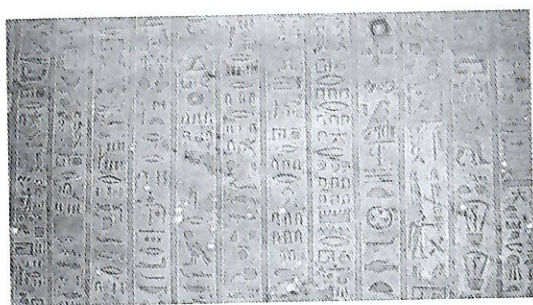
# Math-Jeunes

## La statistique

Dès son apparition sur terre, *Homo Sapiens* a vraisemblablement été amené à effectuer de petits dénombrements. Au fur et à mesure de la diversification de ses activités et de leur complexification, des démarches intellectuelles nouvelles ont été mises en œuvre. Le passage du stade de la chasse et de la cueillette, à celui de l'élevage et de l'agriculture a certainement contribué à accélérer le phénomène. En effet, les hommes se sédentarisent, ils se rassemblent en groupes de plus en plus importants. Le stockage de surplus de nourriture permet la spécialisation du travail. Certains individus fabriquent des outils ou des armes qu'ils échangent contre les vivres produits par les paysans.

Une lente évolution débouche sur la création de sociétés structurées, hiérarchisées. Pour les administrer, les petits dénombrements ne suffisent plus. Il s'agit par exemple de savoir quelle quantité de nourriture est stockée, où et comment elle est entreposée, à quelle rythme elle est consommée et à quelle rythme elle est renouvelée. Si des ennemis se manifestent, ce sont des hommes en âge de porter les armes qu'il convient de dénombrer, puis de répartir en groupes, d'équiper et d'entraîner au combat.

### Un inventaire d'offrandes sur les murs du temple égyptien d'Edfou



Tout groupe humain établi en société est ainsi obligé de réaliser des activités que nous rangeons actuellement sous le terme général de *statistique*<sup>(1)</sup>. C'était vrai dès l'antiquité. Par exemple, les tablettes babyloniennes et les murs des temples égyptiens com-

portent des inventaires de propriétés, d'offrandes, d'impôts...

C'est encore plus vrai à notre époque. Tout état organisé relève de nombreuses données chiffrées qui sont invoquées pour prendre des décisions et élaborer des politiques à suivre. Une petite visite sur le site Web de l'*Institut National de Statistique*, <http://statbel.fgov.be/>, permet de se rendre compte de l'importance quantitative et qualitative des données récoltées. Tous les secteurs de l'activité humaine sont concernés, des plus importants (emploi, commerce, industrie, environnement...) aux plus anecdotiques (fréquence des prénoms des nouveaux-nés...).

Nous sommes donc inondés de nombres et de statistiques. Et bien que l'accord soit général sur le fait que ces statistiques sont indispensables, il serait hasardeux d'affirmer qu'elles sont populaires! N'entend-on pas souvent des phrases du genre *On peut faire dire aux chiffres n'importe quoi! Ces chiffres sont sans doute corrects, mais que peut-on en faire? que signifient-ils vraiment?*

En effet, la statistique ne se contente pas de *décrire* une situation, elle veut aussi *analyser* cette situation. Et c'est là que les vraies difficultés commencent.

En 1999, la SBPMEF publiait un petit fascicule dont nous extrayons le paragraphe suivant :

*[...] une analyse (statistique ou non) de situation se fait toujours dans un contexte et avec un projet : sur un sujet qui est débattu dans la société ou auquel on s'intéresse personnellement, on essaie d'y voir plus clair, pour se faire une opinion mieux établie, pour faire le point sur des idées reçues, pour argumenter ou contredire des prises de position, parce qu'on a une décision à prendre...*

Ce qui est en cause, ce n'est pas tellement la *description* numérique de la réalité — pour autant qu'elle soit claire et complète — que l'*interprétation* de

<sup>(1)</sup> On a souvent fait remarquer que dans « statistique », il y a « état » mais d'aucuns contestent que là soit l'origine du mot « statistique. »

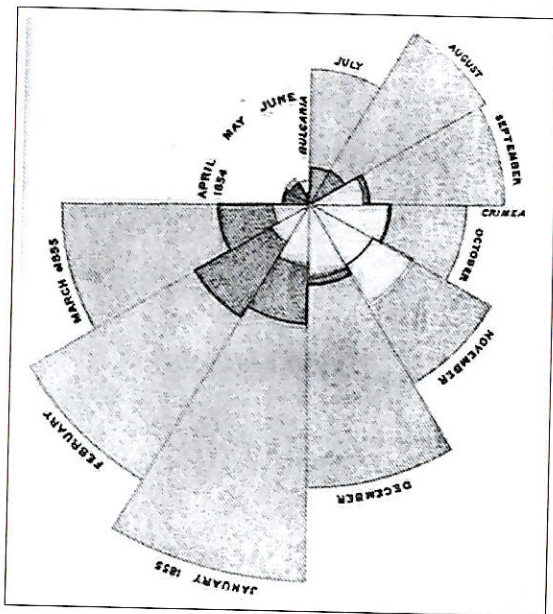


cette description. De fait, dans un conflit, chacune des parties intéressées a son propre projet, souvent incompatible avec celui des autres, et par conséquent chacune a aussi son interprétation des données numériques, en ne retenant souvent que celles qui semblent lui donner raison.

C'est vrai qu'il n'est pas toujours facile d'y voir clair, d'autant plus que bon nombre des problèmes sous-jacents sont complexes et se réduisent mal, sinon pas du tout à une avalanche de chiffres.

Pour compenser l'aridité des tableaux de nombres, il est d'usage de réaliser des représentations graphiques de toutes sortes. Le diagramme suivant est un des premiers graphiques statistiques à avoir été réalisés. Il est dû à l'infirmière anglaise Florence Nightingale (1820-1910) qui l'utilisa pour convaincre les autorités anglaises d'améliorer les conditions hospitalières. Nous sommes ici contraints de le reproduire sans les couleurs originales.

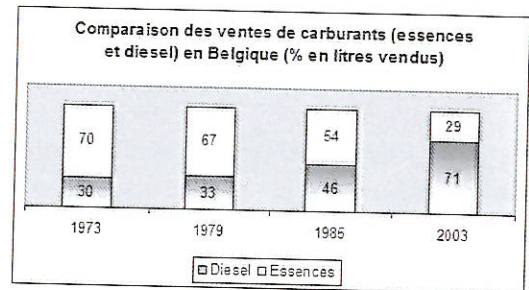
### Un diagramme polaire



Ce diagramme polaire représente les décès survenus durant la guerre de Crimée. Il commence en juillet 1854. Le disque est découpé en douze secteurs correspondant aux douze mois de l'année. Chaque secteur comporte trois parts. L'une, en bleu (la plus à l'extérieur), représente les décès dus à des maladies évitables avec de meilleures conditions de soins et d'hygiène (malnutrition, froid, choléra, dysenterie, typhus, scorbut...). La surface des parts roses (proches du centre) est proportionnelle au nombre de décès dus aux blessures de guerre. En vert (entre les parts roses et les parts bleues), les morts dues à diverses causes. Il apparaît ainsi qu'une meilleure hygiène permettrait de diminuer considérablement le nombre de décès.

(Texte de Hannan Acherki, élève à l'Athénée Royal Gatti de Gamond, que nous remercions vivement pour sa collaboration.)

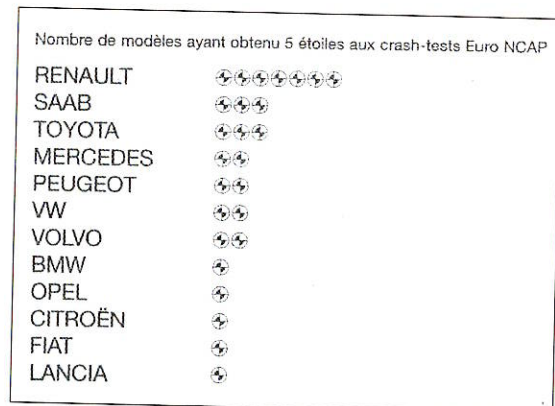
### Un diagramme facile à interpréter



(Source : Institut National de Statistique)

Encore faut-il que la présentation des graphiques ne fausse pas l'interprétation des données.

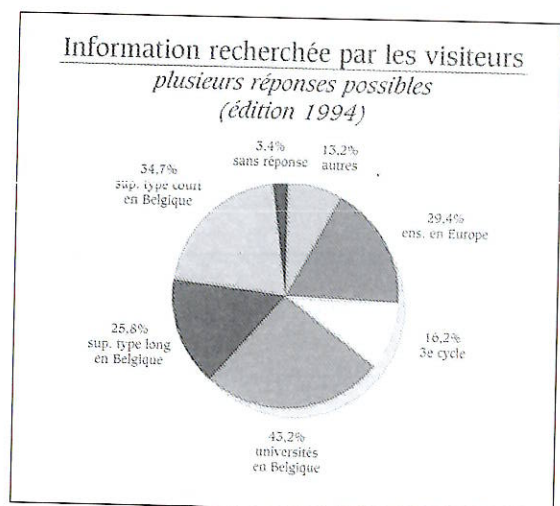
### Une description incomplète



(Source : Publicité dans La Libre, 15-10-04)

Question : COMBIEN DE MODÈLES chaque marque automobile a-t-elle soumis au test ?

### Un diagramme inadapté



(Source : Brochure du Salon européen de l'étudiant, 1995)

Malgré la précaution « plusieurs réponses possibles », ce « camembert » dont le total des



portions est supérieur à 100% a un goût inhabituel!

## Décrire une situation

Un des premiers objectifs de la statistique descriptive est de résumer une situation complexe, comportant de nombreuses données numériques, à l'aide d'un petit nombre de paramètres. Les plus connus sont les paramètres de position (moyenne, médiane, mode) et les paramètres de dispersion (écart-type, moyenne des écarts absolus, étendue). Le choix des paramètres à utiliser pour la description fait déjà partie de l'analyse : selon la situation à analyser, un paramètre convient mieux qu'un autre.

**Quel paramètre de position est le plus adéquat ?** (Les montants mentionnés dans le tableau qui suit sont en euros.)

	Revenu 2002, exercice 2003
Revenu moyen par déclaration	24.621
Région de Bruxelles-Capitale	23.776
Région flamande	25.620
Région wallonne	23.018
Revenu médian	18.914
Région de Bruxelles-Capitale	17.588
Région flamande	19.764
Région wallonne	17.692

(Source : Institut National de Statistique)

Dans le présent fascicule, nous ne nous pencherons pas plus longuement sur les questions liées au choix et à l'usage de ces paramètres. Elles ont été traitées de façon remarquable dans la brochure de la SBP-MEF déjà citée plus haut. Cette brochure est toujours disponible et parfaitement accessible aux lecteurs de *Math-Jeunes*.

(S'adresser au secrétariat de la SBP-Mef, voir p.3 de couverture.)

Statistiques  
Dossier didactique n°6 de la  
SBP-Mef



## Analyser une situation

L'expression « Analyser une situation » peut recouvrir des démarches très différentes. Par exemple :

- Il peut s'agir de rechercher les « facteurs » qui jouent le rôle le plus important dans l'évolution d'un phénomène donné.

Quels sont les facteurs qui influencent le plus les dépenses en matière de soins de santé ?

On pourrait penser

- au vieillissement de la population,
- à l'évolution du prix des médicaments,
- au coût des appareils médicaux,
- ...

- Il peut s'agir de tester une hypothèse

Si on constate qu'une même personne gagne au premier rang au Lotto lors de deux semaines consécutives, l'hypothèse que les tirages du Lotto ne sont pas biaisés est-elle encore acceptable ?

- Il peut s'agir de construire un modèle théorique d'une situation concrète.

Dans ce numéro de *Math-Jeunes* nous vous proposons — en plus des rubriques habituelles — essentiellement des articles présentant précisément des modélisations théoriques de situations concrètes. Comme plusieurs de ces modélisations font appel à des concepts mathématiques relativement élaborés comme les exponentielles et logarithmes, il nous a paru opportun d'insérer les prérequis nécessaires — sans démonstrations ni calculs — aux pages 25 à 27.

Dans *Elle tremble, elle tremble*, G. Noël part des statistiques de tremblements de terre pour élaborer la « loi » de Gutenberg-Richter, qui décrit la fréquence des tremblements de terre en fonction de leur magnitude (sur l'échelle de Richter).

Ensuite, dans *Tout droit dans le nuage*, P. Tilleuil traite d'un problème fréquent : la recherche d'une relation du premier degré entre deux séries appariées de données. Cela s'interprète comme la recherche d'une droite qui passe le plus près possible d'un « nuage » de points non alignés. Encore faut-il savoir ce que signifie « le plus près possible ». Au moins deux interprétations sont possibles... et donc au moins deux droites différentes peuvent être envisagées.

J.-M. Desbonnez nous montre dans *Le calcul statistique d'une aire* comment utiliser un tableur et une simulation pour déterminer approximativement l'aire sous une courbe.

Enfin, l'élaboration de contrats d'assurance fait un usage intensif de statistiques afin d'évaluer les risques que prennent les assureurs et donc les primes que les assurés devront payer. Cette composante importante de l'activité économique d'un pays est abordée par D. Justens qui traite de l'élaboration des *tables de mortalité* à partir de l'observation des décès.



# Edmund Halley

Simone Trompler



E. Halley  
(1656-1742)

Il y a 350 ans naissait l'anglais Edmund Halley. C'était le fils d'un riche fabricant de savon, qui perdit une grosse partie de sa fortune dans l'incendie de Londres, en 1666. Néanmoins, il resta aisé et put donner à son fils la meilleure éducation.

Très tôt, Halley fut un excellent élève et entra au *Queen's College* d'Oxford à dix-sept ans. À l'époque, il était déjà un astronome de qualité, disposant d'une grande collection d'instruments. Il fit d'importantes observations, notamment une occultation de Mars par la Lune en 1676. À cette date, il abandonna ses études pour naviguer vers Sainte Hélène, dans l'hémisphère Sud. Les étoiles de l'hémisphère Nord étaient bien cataloguées, et Halley voulait faire de même pour l'hémisphère Sud. Ce voyage fut financé par son père, mais aussi par le Roi Charles II. Son travail l'amena, entre autres, à améliorer le sextant, à étudier l'océan et l'atmosphère.

Il était dès lors célèbre, et fut élu membre de la *Royal Society* en 1678, devenant ainsi un de ses plus jeunes membres. Il fut un fervent admirateur et ami de Newton et paya de sa poche l'édition des « *Principia Mathematica* » que Newton n'avait pu publier parce que la *Royal Society* manquait de fonds.

Il n'obtint pas la chaire d'astronomie à Oxford, qui était vacante car les autorités craignaient qu'il ne « corrompe la Jeunesse ». En effet, il doutait de l'exactitude scientifique de l'histoire de la Création, dans la Bible !

Il poursuivit néanmoins ses observations et réalisa d'autres travaux. Par exemple, il établit des tables de mortalité (voir dans ce numéro *Histoire et mathématiques des tables de mortalité*, par D. Justens) pour la ville de Breslau. Il s'agissait là d'une des premières tentatives de ce genre.

Table I, by Dr. Halley.

Age	Living	Age	Living	Age	Living	Age	Living	Age	Living	Age	Living
1	1000	16	622	31	523	46	387	61	232	76	78
2	855	17	616	32	515	47	377	62	222	77	68
3	798	18	610	33	507	48	367	63	212	78	58
4	760	19	604	34	499	49	357	64	202	79	49
5	732	20	598	35	490	50	346	65	192	80	41
6	710	21	592	36	481	51	335	66	182	81	34
7	692	22	586	37	472	52	324	67	172	82	28
8	680	23	580	38	463	53	313	68	162	83	23
9	670	24	574	39	454	54	302	69	152	84	19
10	661	25	567	40	445	55	292	70	142	*	*
11	653	26	560	41	436	56	282	71	131		
12	646	27	553	42	427	57	272	72	120		
13	640	28	546	43	417	58	262	73	109		
14	634	29	539	44	407	59	252	74	98		
15	628	30	531	45	397	60	242	75	88		

La table de mortalité de Halley, extraite de l'ouvrage d'Abraham de Moivre, *The doctrine of chances*, (1756).

Il est bien connu que Halley établit que les comètes aperçues en 1305, 1380, 1456, 1531 et 1607 étaient en fait les apparitions successives d'une même comète (appelée depuis la comète de Halley), qui reviendrait tous les 75 ans.

Cette découverte établissait que les comètes peuvent avoir des trajectoires elliptiques, alors que Newton pensait qu'elles étaient toutes paraboliques. Il fit des voyages à bord de navires, déterminant avec précision la longitude des lieux parcourus.

Il étudia aussi l'archéologie, la géophysique, l'histoire de l'astronomie et la solution des équations polynomiales. Ce savant très créatif et polyvalent joua un rôle important dans la communauté scientifique anglaise de son époque.



# Olympiades

Claudine Festraets

## Participer à l'Olympiade !

Durant cette année scolaire, aura lieu la trente et unième édition de l'Olympiade Mathématique Belge (OMB). Voici le calendrier de cette OMB : éliminatoire : le 18 janvier 2006, demi-finale : le 8 mars 2006, finale : le 26 avril 2006, proclamation : le 13 mai 2006.

## Se préparer !

Si tu souhaites te préparer à cette épreuve, tu trouveras ci-dessous quelques exercices posés dans le passé. Les solutions se trouvent à la page 32. Ne les regarde pas tout de suite. Si tu désires peaufiner ta préparation et évaluer ton potentiel olympique, tu peux acquérir le tome 5 des OMB, il contient toutes les questions posées de 1999 à 2002, demande à ton professeur comment on peut se le procurer.

## S'exercer !

### 1. Trapèze (demi-finale - 1990)

Les angles consécutifs d'un trapèze forment une progression arithmétique. Si le plus petit angle mesure  $75^\circ$ , le plus grand mesure

- (A)  $95^\circ$  (B)  $100^\circ$  (C)  $105^\circ$  (D)  $110^\circ$  (E)  $115^\circ$

### 2. Quelques calculs (éliminatoire - 1992)

$$\frac{1}{2} + \left(\frac{1}{3} + \frac{2}{3}\right) + \left(\frac{1}{4} + \frac{2}{4} + \frac{3}{4}\right) + \dots + \left(\frac{1}{100} + \frac{2}{100} + \dots + \frac{99}{100}\right) =$$

- (A) 105 (B) 245 (C) 2475 (D) 3215 (E) 4950

### 3. Second degré (éliminatoire - 1989)

Si  $x^2 - 5x + p \neq 0$  pour tout nombre réel  $x$ , une seule des valeurs suivantes peut être prise par  $p$ . Laquelle ?

- (A) 0 (B) 4 (C) 6 (D) -6 (E) 8

### 4. Coordonnées (demi-finale - 1989)

Si  $a, b > 0$  et si le triangle limité dans le premier quadrant par les axes de coordonnées et la droite  $ax + by = 6$  est d'aire 6, alors  $ab =$

- (A) 3 (B) 6 (C) 12 (D) 108 (E) 432

### 5. Fonction (demi-finale - 1990)

Soit  $f$  la fonction définie par  $f(x) = ax^2 - \sqrt{2}$  pour un certain réel  $a$  strictement positif. Si  $f(f(\sqrt{2})) = -\sqrt{2}$ , alors  $a =$

- (A)  $\frac{2-\sqrt{2}}{2}$  (B)  $\frac{1}{2}$  (C)  $2 - \sqrt{2}$  (D)  $\frac{\sqrt{2}}{2}$  (E)  $\frac{2+\sqrt{2}}{2}$

### 6. Système (éliminatoire - 1988)

Les réels  $u, v$  sont tels que  $\begin{cases} u + v = 2 \\ u^2 + v^2 = 3 \end{cases}$

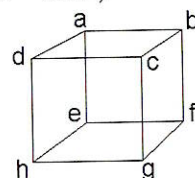
Que vaut  $u^3 + v^3$  ?

- (A)  $\frac{7}{2}\sqrt{2}$  (B) 5 (C)  $3\sqrt{3}$  (D) 6 (E) 8

### 7. Cube et vecteurs (éliminatoire - 1992)

Dans la figure,  $a, b, c, d, e, f, g, h$  sont les sommets d'un cube.

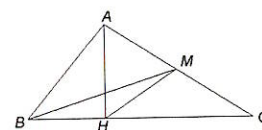
Alors  $\vec{bh} - \vec{gf} + \vec{ca} + \vec{hd} - \vec{bd} + \vec{cg} =$



- (A)  $\vec{ah}$  (B)  $\vec{bh}$  (C)  $\vec{ch}$  (D)  $\vec{df}$  (E)  $\vec{eb}$

### 8. Mesure d'angle (demi-finale - 1989)

Les angles du triangle  $ABC$  valent  $\hat{A} = 100^\circ$ ,  $\hat{B} = 50^\circ$ ,  $\hat{C} = 30^\circ$ . Si  $AH$  est une hauteur et  $BM$  une médiane, l'angle  $\widehat{MHC}$  mesure



- (A)  $15^\circ$  (B)  $22,5^\circ$  (C)  $30^\circ$  (D)  $40^\circ$  (E)  $45^\circ$

### 9. Chaussures (éliminatoire - 1989)

Un tiroir contient trois chaussures pour pied gauche et deux chaussures pour pied droit. la probabilité qu'en prenant, au hasard et simultanément, deux chaussures dans ce tiroir on en ait une pour pied gauche et une pour pied droit vaut

- (A)  $\frac{6}{10}$  (B)  $\frac{2}{5}$  (C)  $\frac{1}{3}$  (D)  $\frac{5}{12}$  (E)  $\frac{7}{12}$

### 10. Fractions (demi-finale - 1988)

Si  $\frac{b}{a} = 2$  et  $\frac{c}{b} = 3$ , que vaut le rapport  $\frac{a+b}{b+c}$  ?

- (A)  $\frac{1}{3}$  (B)  $\frac{3}{8}$  (C)  $\frac{3}{5}$  (D)  $\frac{2}{3}$  (E)  $\frac{3}{4}$



# Elle tremble, elle tremble!

Guy Noël

## Pour en savoir plus

[1] Per Bak, *Quand la nature s'organise. Avalanches, tremblements de terre et autres cataclysmes*, Flammarion, 1999.

### Sites internet :

[2] Site de l'U.S Geological Survey :

[neic.usgs.gov/neis/eqlists/eqstats.html](http://neic.usgs.gov/neis/eqlists/eqstats.html)

[3] Site du Nevada Seismological Laboratory :

[www.seismo.unr.edu/index.html](http://www.seismo.unr.edu/index.html)

Le tsunami du 26 décembre 2004 est encore dans toutes les mémoires ! Faut-il rappeler que cette catastrophe a été provoquée d'abord par un tremblement de terre de magnitude 9 sur l'échelle de Richter ? Heureusement, pensez-vous peut-être, les tremblements de terre sont relativement rares. En êtes-vous si sûr ?

Cherchons des renseignements sur Internet. Sur le site de l'*United States Geological Survey*, nous lisons (en anglais) :

*L'USGS estime que plusieurs millions de tremblements de terre surviennent chaque année dans le monde. Beaucoup ne sont pas détectés car ils touchent des régions éloignées ou sont de très faible amplitude. Le NEIC (National Earthquake Information Center) localise environ cinquante tremblements de terre par jour, ou environ vingt mille par an.*

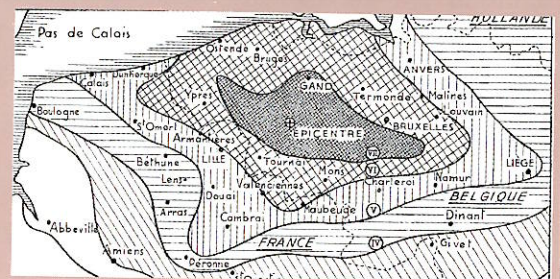
Par ailleurs, on pourrait considérer comme normal que les séismes de faible magnitude soient plus fréquents que les autres. Peut-on étudier comment cette fréquence varie ?

Le tableau suivant (extrait de [2]), indique, en fonction de leur magnitude, le nombre de tremblements de terre survenus au cours des années 2000 à 2004, en un point quelconque du globe terrestre :

Magnitude	2000	2001	2002	2003	2004
8,0 à 9,9	1	1	0	1	2
7,0 à 7,9	14	15	13	14	13
6,0 à 6,9	158	126	130	140	139
5,0 à 5,9	1345	1243	1218	1203	1430
4,0 à 4,9	8045	8084	8584	8462	10975
3,0 à 3,9	4784	6151	7005	7624	7826
2,0 à 2,9	3758	4162	6419	7727	6297
1,0 à 1,9	1026	944	1137	2506	1340
0,1 à 0,9	5	1	10	134	103
Sans Magnitude	3120	2938	2937	3608	3076
Total	22256	23534	27454	31419	31201

*Le plus important tremblement de terre enregistré en Belgique au XX<sup>e</sup> siècle.*

Le 11 juin 1938, un séisme de magnitude 5,6 frappe la Belgique. Son épicentre est situé près de Renaix. À l'observatoire d'Uccle, les sismographes sont mis hors service ! Les dégâts vont jusqu'en France.



Dessin extrait de *La Terre, ma patrie*, par Pierre Rousseau, Ed. Fayard, 1947.



Comme tout tableau de données numériques, celui-ci doit être lu d'un œil critique. Ce qui nous intéresse, c'est d'étudier comment le nombre de tremblements de terre varie en fonction de la magnitude. De ce point de vue, la ligne recensant les séismes — sans doute trop faibles — auxquels il n'a pas été possible d'attribuer une magnitude ne nous sert à rien.

Nous constatons que d'année en année le nombre total de tremblements de terre augmente sensiblement. En cinq ans, on passe en effet de 22 256 séismes à 31 201 ! Tremble-t-elle donc de plus en plus ?

En examinant plus attentivement encore le tableau, nous observons surtout une forte augmentation des séismes de faible amplitude. Par contre, le nombre de séismes d'amplitude au moins 4, reste relativement stable d'une année à l'autre.

Une explication de ce phénomène peut se trouver dans l'évolution du nombre de stations sismographiques. Le site internet mentionné ci-dessus nous apprend qu'en 1931, 350 stations de ce genre existaient sur terre. Aujourd'hui on en dénombre plus de 8000. De plus la sensibilité des appareils augmente ! Ainsi, on peut considérer que les séismes suffisamment importants ont presque toujours été détectés, mais que le nombre de petits séismes *détectés* augmente d'année en année. Seules les cinq premières lignes du tableau doivent donc être considérées comme significatives pour l'étude que nous voulons faire.

Dressons un nouveau tableau du nombre moyen annuel de séismes pour les années 2000–2004, en fonction de la magnitude :

Magnitude	4,0 à 4,9	5,0 à 5,9	6,0 à 6,9	7,0 à 7,9	8,0 à 9,9
Fréquence moyenne	8830	1287,8	138,6	13,8	1

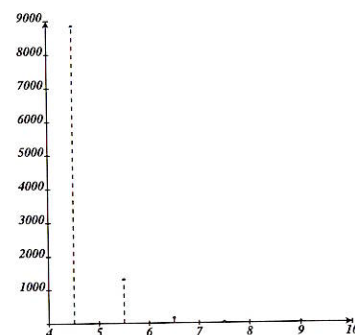


Figure 1

Effectuons une représentation graphique « en bâtonnets » (Figure 1) : Sur la figure 1, chaque classe du tableau ci-dessus est représentée par un point dont l'abscisse est au centre de la classe et dont l'ordonnée est la fréquence de la classe.

Reconnaissons qu'un tel diagramme est mal adapté à la situation du fait que la variable considérée (ici la fréquence de la classe) a un très grand intervalle de variation (ici de 1 à 8830 !). Dans un cas pareil, on a souvent intérêt à porter en ordonnée, non la variable elle-même, mais son logarithme (lisez éventuellement la page 26). On dit qu'on réalise un diagramme « logarithmique ». Dans le cas présent, ce logarithme varie de  $0 = \log_{10} 1$  à  $3,946 = \log_{10} 8830 \dots$  ce qui détermine un intervalle beaucoup plus petit.

Voici donc le tableau des logarithmes des fréquences (arrondis à 3 décimales) et le diagramme correspondant (figure 2).

Magnitude	4,0 à 4,9	5,0 à 5,9	6,0 à 6,9	7,0 à 7,9	8,0 à 9,9
Logarithme de la fréquence moyenne	3,946	3,110	2,142	1,140	0

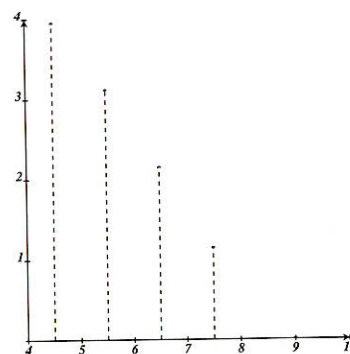


Figure 2

Voilà un diagramme beaucoup plus parlant ! Et qui nous interpelle : n'avez-vous pas l'impression que les sommets des bâtonnets sont alignés ? Pour apprécier ce phénomène, calculons la droite qui approche au mieux notre « nuage » de cinq points.



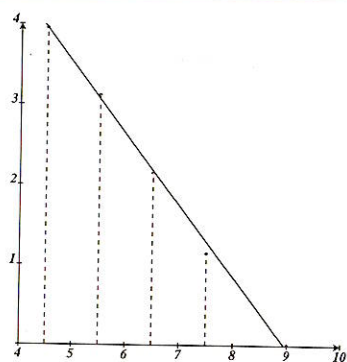


Figure 3

En cherchant la droite « des moindres carrés », c'est-à-dire celle qui minimise l'erreur quadratique, (voir dans ce numéro l'article *Tout droit dans le nuage*), on obtient l'équation

$$y = -0,89x + 7,96$$

La figure 3 est obtenue en superposant à la figure 2 la droite ayant cette équation.

Ainsi, dans notre cas particulier au moins, nous constatons qu'en effet il est raisonnable de considérer que le logarithme de la fréquence des séismes est une fonction affine de la magnitude<sup>(1)</sup>. Autrement dit, nous considérons que la formule

$$\log_{10} f = -0,89 \cdot m + 7,96$$

constitue une *modélisation acceptable* de la relation entre la magnitude  $m$  et le logarithme de la fréquence moyenne annuelle des séismes de magnitude  $m$ .

Cette constatation est en fait tout à fait générale et connue sous le nom de **Loi de Gutenberg-Richter**. Elle est valable quelle que soit la période de temps sur laquelle les moyennes sont calculées, et même si on limite le relevé des séismes à ceux qui surviennent dans une région limitée du globe terrestre. Bien entendu la pente et l'ordonnée à l'origine peuvent varier, mais l'approximation par une droite est toujours très bonne.

À titre d'exemple, voici les données concernant la Californie (région très affectée par des tremblements de terre) et les années 2000 à 2004 :

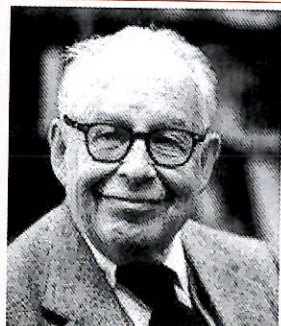
Magnitude	4,0 à 4,9	5,0 à 5,9	6,0 à 6,9	7,0 à 7,9	8,0 à 9,9
Fréquence moyenne	390,2	49,6	5,8	0,8	0

Procédez au même travail que ci-dessus, en n'oubliant pas que la fonction logarithme n'est pas définie en 0. Votre nuage n'aura donc que 4 points au lieu de 5. Calculez la droite qui minimise l'erreur quadratique. Vous devez trouver l'équation

$$\log_{10} f = -0,900m + 6,636$$

À partir du moment où on modélise un phénomène naturel à l'aide d'une formule mathématique, il devient possible de faire des prédictions. Ainsi, à l'occasion du *tsunami* de décembre 2004, les journaux ont abondamment rappelé que la Californie vit dans l'attente d'un « big quake » : un séisme très violent qui sera provoqué par le déplacement des plaques tectoniques et qui aurait deux « chances » sur trois de se produire avant 2020. Un tel phénomène semble *a priori* ne rien devoir au hasard. Et pourtant la loi de Gutenberg-Richter nous fournit une information d'ordre statistique.

<sup>(1)</sup> Des techniques statistiques plus élaborées (test du «  $\chi^2$  ») permettent d'évaluer le risque de se tromper en admettant cette hypothèse, mais elles débordent nettement du cadre de cet article.



Charles Richter  
1900–1985

Dès 1931, l'américain Richter collabora avec son collègue du California Institute of Technology, Beno Gutenberg, pour étudier les phénomènes sismiques. C'est en 1935 qu'il élaborait l'échelle qui porte son nom, utilisée depuis pour mesurer la magnitude des séismes.



Beno Gutenberg  
1889–1960



Par exemple d'après le modèle  $\log_{10} f = -0,900m + 6,636$ , valable pour la Californie, pour  $m = 8,5$ , on trouve  $\log_{10} f = -1,014$ , donc  $f = 10^{-1,014} = 0,0968 \dots$  Il peut paraître bizarre de dire qu'il se produit en Californie 0,0968 séisme(s) de magnitude 8,5 par an ... sauf si on se rappelle qu'il s'agit d'une moyenne. Comme  $\frac{1}{0,0968} = 10,33$ , un tremblement de terre de cette importance apparaît en moyenne une fois tous les dix ou onze ans.

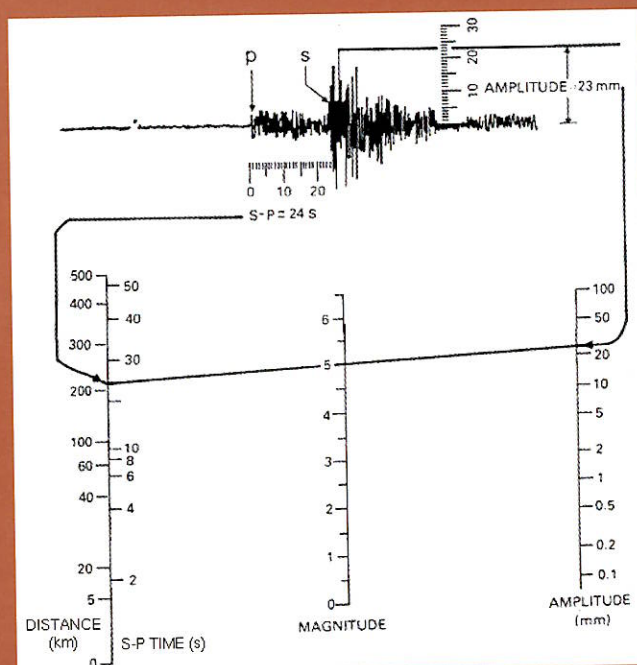
## L'échelle de Richter

En 1935, Richter définit l'amplitude d'un tremblement de terre comme étant le logarithme de l'amplitude maximum des ondes sismiques. De façon plus précise, il utilise la formule suivante :

$$m = \log_{10} a + 3 \log_{10}(8\Delta) - 2,92$$

où  $a$  est l'amplitude maximum des oscillations du stylet du sismographe, mesurée en mm et où  $\Delta$  est la distance du sismographe à l'épicentre du séisme, mesurée par le temps mis par les ondes sismiques pour atteindre le sismographe.

Pour déterminer aisément la magnitude d'un tremblement de terre, Richter utilise un « nomogramme » :



Dessin extrait de [3]

La définition donnée par Richter caractérise la magnitude d'un séisme par le déplacement de la croûte terrestre. Mais ce qui provoque les destructions, c'est l'énergie libérée. Aussi, en 1956, Richter convertit cette définition en une expression faisant intervenir l'énergie libérée par le séisme. Avec Beno Gutenberg, il montre qu'un tremblement de terre de magnitude 0 libère une énergie de l'ordre de  $10^4$  joules, cependant que la magnitude 9 correspond à une énergie de l'ordre de  $10^{17}$  joules. L'échelle de Richter étant logarithmique, l'énergie libérée par un séisme de magnitude  $m + 1$  est donc approximativement égale au produit de celle libérée par un séisme de magnitude  $m$  par la racine neuvième de  $\frac{10^{17}}{10^4}$ , soit environ 30. Un séisme de magnitude 8 est par conséquent environ 900 plus « destructeur » qu'un autre de magnitude 6 !



# Tout droit dans le nuage!

Philippe Tilleuil

## Un exemple concret

Voici un exemple concret d'une situation où il peut être intéressant de mettre en évidence une relation entre deux variables.

Une étude médicale à propos du tabagisme vise à comparer la consommation annuelle de cigarettes par tête, notée  $C$ , au taux de mortalité due au cancer du poumon par 100 000 habitants, noté  $M$ . Menée dans 6 états américains, cette étude a fourni les résultats ci-dessous.

Etat	$C$	$M$
Delaware	3400	24
Indiana	2600	20
Iowa	2200	17
Montana	2400	19
New Jersey	2900	26
Washington	2100	20

## Une représentation graphique

Il est toujours plus parlant de représenter graphiquement des résultats mettant en jeu deux variables. On appelle *nuage de points* un ensemble de points dont les coordonnées dans un repère cartésien sont les couples correspondants d'un tableau statistique à deux variables. Dans le cas qui nous occupe, ce sont les points dont les coordonnées sont les couples  $(C; M)$  du tableau ci-dessus. Ils sont représentés dans la figure 1 ci-dessous, à gauche.

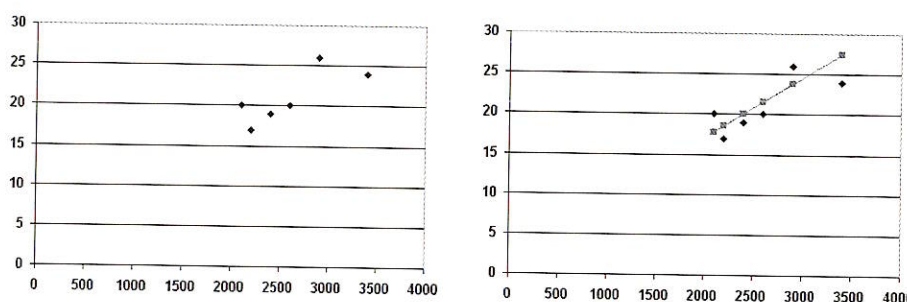


Figure 1 : Le nuage de points pour l'étude sur le tabagisme. Le graphique de la fonction  $f(x) = 0,0075x + 2$  pourrait « représenter » ce nuage.

Il semble probable qu'une augmentation de la consommation  $C$  entraîne une augmentation de la mortalité  $M$ . S'il existait une relation entre les abscisses et les ordonnées des points du nuage, nous pourrions l'utiliser — dans certaines limites — pour *prédire* le taux de mortalité  $M$  par cancer du poumon à partir de la consommation  $C$  de tabac. Nous allons donc rechercher une telle relation.

La relation la plus simple possible entre ces deux variables est du premier degré et s'écrit donc  $M = f(C) = aC + b$ . Mais il ne faut certainement pas s'attendre à ce que cette relation soit *exacte* : le cancer du poumon n'est pas exclusivement dû au tabagisme. Il n'y a donc aucune raison pour que les points du nuage soient strictement alignés. Tout ce qu'on peut espérer, c'est que le graphique de la fonction  $f(C) = aC + b$  représente « au mieux » le nuage de points considéré. C'est ce qui semble être le cas dans la figure 1 ci-dessus, à droite.

Mais il peut y avoir beaucoup de fonctions du premier degré qui pourraient convenir... Pourquoi l'une d'entre elles serait-elle mieux placée qu'une autre vis-à-vis du nuage? Que veut dire « au mieux »?

La solution consiste à définir *a priori* une *mesure d'erreur*, qui compare les valeurs données aux valeurs calculées par la fonction, et de convenir qu'une fonction du premier degré représente mieux le nuage qu'une autre si sa mesure d'erreur est plus petite que celle de l'autre. En ce sens, la *meilleure* fonction est celle qui rendra la fonction d'erreur aussi petite que possible, ... si c'est possible.

Les fichiers EXCEL suivants sont disponibles sur le site Internet de la SBPMef (adresse : [www.sbp.m.be](http://www.sbp.m.be)) :

Comparaison.xls  
ErreurTotale(Général).xls  
ErreurTotale(Barycentre).xls  
ErreurAbsolue.xls  
ErreurQuadratique.xls



## Un autre exemple, encore plus simple

L'exemple du tabagisme est encore un peu compliqué puisqu'il comporte déjà 6 points. Limitons-nous à une question beaucoup plus simple. Comment trouver une fonction du premier degré qui approche au mieux un nuage qui comporte seulement 3 points ? Cet exemple-là est certainement le plus simple de tous, puisque dans un nuage de 2

Nom	$x$	$y$
A :	-5	-6
B :	8	-4
C :	5	8

points, il n'y aurait tout simplement pas de problème du tout ! Pour fixer les idées, considérons désormais les 3 points dont les coordonnées sont décrites dans le tableau ci-contre, et qui sont représentés dans le graphique à côté.

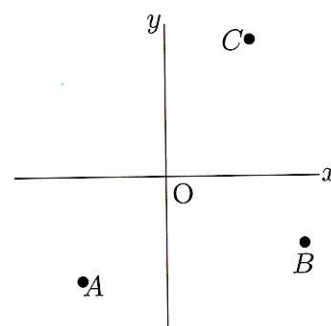


Figure 2 : Les positions des trois points dans le plan.

Nous voulons donc déterminer des valeurs de  $a$  et  $b$  de telle sorte que le graphique de la fonction  $f(x) = ax + b$  représente au mieux ces 3 points. Pour cela, nous allons définir et étudier quelques mesures d'erreur.

### Trois mesures d'erreur

Notons  $(x_A; y_A)$ ,  $(x_B; y_B)$  et  $(x_C; y_C)$  les coordonnées des points A, B et C. À partir des trois nombres  $y_A - f(x_A)$ ,  $y_B - f(x_B)$  et  $y_C - f(x_C)$ , nous allons définir trois mesures d'erreur, relativement simples et naturelles : la *valeur absolue de l'erreur totale*, l'*erreur absolue*, et l'*erreur quadratique*. Pour chacune de ces trois mesures d'erreur, nous allons chercher quelles sont les fonctions du premier degré  $f(x) = ax + b$  qui la rendent minimum.

### Ne confondons pas !

Notre problème n'est pas de trouver la droite qui approche le mieux le nuage, mais bien la fonction du premier degré  $f(x) = ax + b$  dont le graphique approche le mieux le nuage.

Pour mesurer l'erreur commise en approchant les ordonnées  $y_i$  des points du nuage par les valeurs  $f(x_i)$  de la fonction  $f$ , nous ne pouvons donc utiliser que les différences  $y_i - f(x_i)$ . Autrement dit, les distances des points du nuage à la droite qui représente la fonction  $f(x) = ax + b$  se mesurent à la verticale et non perpendiculairement !

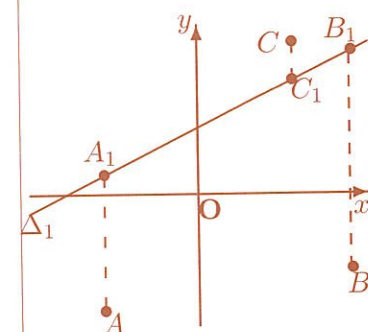


Figure 3

- L'*erreur totale* est donnée par la formule :

$$(y_A - f(x_A)) + (y_B - f(x_B)) + (y_C - f(x_C))$$

Comme une erreur est normalement un nombre positif, nous considérons sa valeur absolue :

$$E_t = |(y_A - f(x_A)) + (y_B - f(x_B)) + (y_C - f(x_C))|$$

C'est la première mesure d'erreur à laquelle on peut penser : elle est calculée à partir de la somme des erreurs « individuelles ».

- L'*erreur absolue* :

$$E_a = |y_A - f(x_A)| + |y_B - f(x_B)| + |y_C - f(x_C)|$$

C'est quasiment la même construction que celle de la valeur absolue de l'erreur totale. Mais une erreur négative ne risque pas de compenser une erreur positive.

- L'*erreur quadratique* :

$$E_q = (y_A - f(x_A))^2 + (y_B - f(x_B))^2 + (y_C - f(x_C))^2$$

C'est peut-être la moins naturelle des mesures d'erreur. A part l'élévation au carré, c'est encore une fois quasiment la même construction que celle des erreurs précédentes. Elle est aussi, nécessairement positive par définition. Par ailleurs, nous verrons plus loin une raison statistique pour laquelle cette erreur est intéressante.



## Une étude expérimentale

On construit une fenêtre graphique dans laquelle sont représentés les trois points  $A$ ,  $B$  et  $C$ , ainsi que le graphique de la fonction  $f(x) = ax + b$ ; on y associe un calcul de l'une ou l'autre erreur, au choix; enfin, on fait varier les valeurs des paramètres  $a$  et  $b$  par exemple de  $-10$  à  $+10$  à  $0,001$  près à l'aide de barres de défilement.

Quand tout cela est bien au point, il reste alors à définir une bonne manière de faire varier les valeurs de  $a$  et de  $b$  pour diminuer autant que possible la mesure d'erreur considérée. C'est le plus gai de l'affaire, et c'est en bricolant une stratégie de minimisation qu'on apprend le plus de choses...

Des fichiers EXCEL® correspondant aux différentes mesures d'erreur peuvent être téléchargés sur le site de la SBPMef. Ils illustrent le problème du tabagisme.

Il s'agit de comprendre comment on peut minimiser ces erreurs. Avant d'établir des résultats théoriques, il vaut mieux faire quelques *expériences* de minimisation. Un tableur, tel EXCEL®, est l'outil rêvé pour ce genre d'activités. Pour l'essentiel, on l'utilise comme suit.

Essayons donc, avec EXCEL®, de minimiser  $E_t$ . (Fichier ErreurTotale(Général).xls)

Après quelques essais, on se rend compte que ça ne va pas (cfr. la figure 4) : dans toutes les directions, il y a une droite telle que

$$\begin{aligned} (y_A - f(x_A)) + \\ (y_B - f(x_B)) + \\ (y_C - f(x_C)) &= 0 \end{aligned}$$

et donc pour laquelle  $E_t = 0$

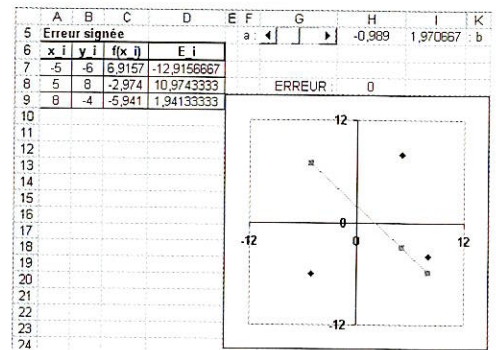


Figure 4 : Quand on s'y prend bien, l'erreur totale peut toujours être nulle !

Plus précisément, il suffit que la droite passe par ce qu'on appelle le *point moyen* du nuage, c'est-à-dire le point dont les coordonnées sont respectivement la moyenne des abscisses et la moyenne des ordonnées.

Dans le cas de notre exemple, le point moyen du nuage a comme coordonnées :  $(\frac{8}{3}; -\frac{2}{3})$ . Quelle est l'équation d'une droite passant par ce point ? On l'obtient immédiatement :

$$y - \left(-\frac{2}{3}\right) = a \left(x - \frac{8}{3}\right) \text{ c'est-à-dire : } y = ax - \frac{8a + 2}{3}$$

Une légère adaptation du fichier EXCEL® (car maintenant  $b$  est une fonction de  $a$ ) permet de se convaincre qu'effectivement, toutes les droites de ce type annulent l'erreur totale. (voir le fichier ErreurTotale(Barycentre).xls)

Pourquoi en est-il ainsi ? C'est assez facile à comprendre. Notons  $x_M = \frac{x_A + x_B + x_C}{3}$  et  $y_M = \frac{y_A + y_B + y_C}{3}$  les coordonnées du point moyen du nuage, l'équation d'une droite passant par ce point moyen s'écrit donc :

$$f(x) = y_M + a(x - x_M)$$

où  $a$  est un nombre réel *quelconque*. L'erreur signée correspondante se calcule alors comme suit :

$$\begin{aligned} E_s &= y_A - (y_M + a(x_A - x_M)) + y_B - (y_M + a(x_B - x_M)) + y_C - (y_M + a(x_C - x_M)) \\ &= y_A + y_B + y_C - 3y_M - a(x_A + x_B + x_C - 3x_M) \\ &= 3 \left\{ \frac{y_A + y_B + y_C}{3} - y_M - a \left( \frac{x_A + x_B + x_C}{3} - x_M \right) \right\} \\ &= 0 \end{aligned}$$

C'est bien ce qu'on voulait établir, et ce n'est quasiment qu'une reformulation de la définition de point moyen du nuage, c'est-à-dire de la définition de moyenne !



L'erreur totale étant hors-jeu, attaquons-nous aux deux autres. Sur la figure 5, on a superposé les résultats de la minimisation pour l'erreur absolue et l'erreur quadratique.

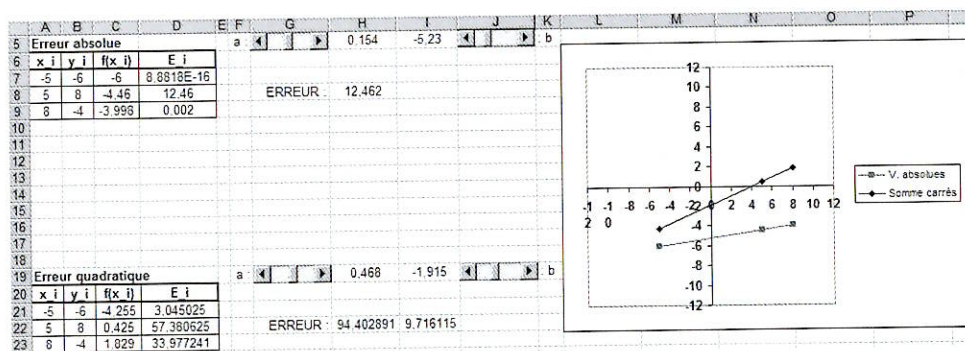


Figure 5 : Le fichier Comparaison.xls permet de comparer deux mesures d'erreurs.

A nouveau, beaucoup d'observations significatives peuvent être faites. Retenons celles-ci :

- La droite qui minimise l'erreur absolue semble passer par deux points du nuage.
- La droite construite sur base de l'erreur quadratique semble passer très près du point moyen du nuage.

## Une propriété de l'erreur absolue

Pourquoi la droite qui minimise l'erreur absolue passe-t-elle toujours par deux points parmi les trois ?

L'idée du raisonnement peut se découvrir grâce à EXCEL®. A direction fixée, quand on ne fait varier que le paramètre  $b$ , on constate que l'erreur diminue quand la droite s'approche d'un des trois points, puis augmente quand elle s'en écarte. Elle passe donc par un minimum (local) lorsque la droite (mobile) contient un des trois points. Et lorsque la droite contient un des trois points, alors, en ne faisant varier que le paramètre  $a$  (c'est-à-dire en n'effectuant plus que des rotations), le phénomène est semblable : l'erreur est minimale lorsque la droite (mobile) passe de plus par un des deux points restants.

Il n'est plus très difficile de formaliser tout cela. On a, (figure 6) :

$$\begin{aligned}
 E_a(\Delta_1) &= |AA_1| + |BB_1| + |CC_1| \\
 &= |AA_2| + |A_2A_1| + |BB_2| + |B_2B_1| + |CC_2| - |C_2C_1| \\
 &= E_a(\Delta_2) + \delta \geq E_a(\Delta_2)
 \end{aligned}$$

car  $\delta = |A_1A_2| + |B_1B_2| - |C_1C_2| \geq 0$ . En adaptant ce raisonnement aux différentes positions possibles de la droite par rapport à  $A$ ,  $B$  et  $C$ , on démontre donc que l'erreur est (localement) minimale lorsque la droite (mobile) contient un au moins des trois points.

D'autre part, on a aussi suivant la figure 7 :

$$\begin{aligned}
 E_a(\Delta_3) &= |BB_3| + |CC_3| \\
 &= |BB_4| + |B_3B_4| + |CC_4| - |C_3C_4| \\
 &= E_a(\Delta_4) + \varepsilon \geq E_a(\Delta_4)
 \end{aligned}$$

car  $\varepsilon = |B_3B_4| - |C_3C_4| \geq 0$  (grâce au théorème de Thalès, puisque  $x_C \in [x_A; x_B]$ ). En adaptant encore une fois ce raisonnement aux différentes positions possibles de la droite par rapport à  $B$  et  $C$ , on démontre donc que l'erreur est (localement) minimale lorsque la droite (mobile) passe par un des deux points restants.

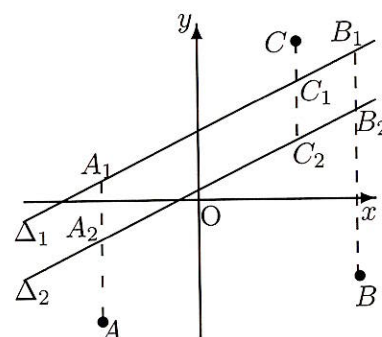


Figure 6 : Par translation, la droite qui minimise l'erreur absolue doit passer par un point du nuage.

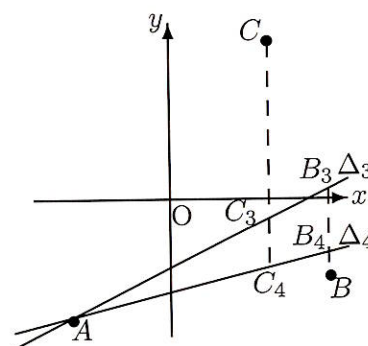


Figure 7 : Par rotation, la droite qui minimise l'erreur absolue doit passer par deux des trois points du nuage.



Il y a donc trois possibilités de minimum (local) pour l'erreur absolue : la droite passant par  $A$  et  $B$ , la droite passant par  $B$  et  $C$ , et la droite passant par  $A$  et  $C$ . En comparant les valeurs de l'erreur pour chaque cas, on détermine immédiatement la droite qui donne lieu au minimum global. Ici, il s'agira de la droite passant par  $A$  et  $B$ .

### Une propriété de l'erreur quadratique

Pour l'erreur quadratique, il y a moyen de réaliser sans trop de peine tous les calculs utiles. Avec les données dont on dispose, l'erreur quadratique s'écrit par définition :

$$\begin{aligned} E_q &= (-6 + 5a - b)^2 + (8 - 5a - b)^2 + (-4 - 8a - b)^2 \\ &= 114a^2 + 16ab + 3b^2 - 76a + 4b + 116 \end{aligned}$$

On peut considérer cette expression comme un trinôme du second degré en $a$ :	Pareillement, on peut considérer la même expression comme un trinôme du second degré en $b$ :
$E_q = 114a^2 + 4a(4b - 19) + 3b^2 + 4b + 116$	$E_q = 3b^2 + 4b(4a + 1) + 114a^2 - 76a + 116$
Ce trinôme a un minimum, puisque le coefficient de $a^2$ est positif, et il est situé sur	Et ce trinôme a lui aussi un minimum, puisque le coefficient de $b^2$ est positif, et il est situé sur
$a = -\frac{4(4b - 19)}{2 \cdot 114} = -\frac{4b - 19}{57} \quad (1)$	$b = -\frac{4(4a + 1)}{2 \cdot 3} = -\frac{2(4a + 1)}{3} \quad (2)$

Les équations (1) et (2) forment un système de deux équations du premier degré à deux inconnues :

$$\begin{cases} 57a + 4b = 19 \\ 8a + 3b = -2 \end{cases} \text{ dont les solutions sont : } \begin{cases} a = \frac{1235}{2641} = 0,467625\dots \\ b = -\frac{266}{139} = -1,913669\dots \end{cases}, \text{ en parfait accord avec les résultats}$$

obtenus grâce à EXCEL<sup>®</sup> (cfr. le bas de la figure 5).

La droite qui minimise l'erreur quadratique (on l'appelle aussi « droite d'approximation au sens des moindres carrés ») est donc la droite d'équation  $y = \frac{1235}{2641}x - \frac{266}{139}$ .

La deuxième équation du système ci-dessus, à savoir  $8a + 3b = -2$  peut s'écrire  $a \cdot \frac{8}{3} + b = -\frac{2}{3}$ , ce qui s'interprète immédiatement : le point moyen du nuage, à savoir le point de coordonnées  $(\frac{8}{3}; -\frac{2}{3})$ , est effectivement un point de la droite  $f(x) = ax + b$ .

Tous ces calculs pourraient évidemment être repris avec des notations plus générales (c'est-à-dire  $(x_A; y_A)$  au lieu de  $(-5; -6)$ , etc...), mais cela ne changerait rien aux idées qui viennent d'être décrites !

### Quelques conclusions ... et prolongements

Résumons-nous ! Nous avons vu que l'erreur totale, ou sa valeur absolue, ne peuvent servir à rien du tout ! Très concrètement, cela signifie qu'il existe des mesures d'erreur avec lesquelles tout le monde pourrait défendre n'importe quoi, en ayant l'air (faussement !) savant.

Il ne reste donc que l'erreur absolue et l'erreur quadratique. Comme signalé plus haut, il y a des raisons *a priori* pour considérer de telles définitions d'erreur. En bref, voici de quoi il s'agit.

On considère un ensemble  $\mathcal{R}$  de  $n$  nombres  $(x_1, x_2, \dots, x_n)$  ordonnés par ordre de grandeur croissante.

- Ce qu'on appelle en statistique la *médiane* de l'ensemble  $\mathcal{R}$  est la valeur de la variable  $X$  pour laquelle la fonction

$$M(X) = \frac{|X - x_1| + |X - x_2| + \dots + |X - x_n|}{n}$$

est minimale. On montre assez facilement que si le nombre d'éléments  $n = 2m + 1$  est impair, alors la médiane de l'ensemble  $\mathcal{R}$  égale  $x_{m+1}$ , tandis que si  $n = 2m$  est pair, alors tous les points de l'intervalle  $[x_m; x_{m+1}]$  conviennent. Dans ce dernier cas, on parle plutôt d'intervalle médian de l'ensemble  $\mathcal{R}$ .



• D'autre part, ce qu'on appelle en statistique la *moyenne* de cet ensemble  $\mathcal{R}$  est la valeur de la variable  $X$  pour laquelle la fonction

$$Q(X) = \frac{(X - x_1)^2 + (X - x_2)^2 + \dots + (X - x_n)^2}{n}$$

est minimale. On montre sans trop de peine que cette valeur est la *moyenne au sens usuel* de ces  $n$  nombres :

$$\frac{x_1 + x_2 + \dots + x_n}{n}$$

En particulier, l'ordre imposé à l'ensemble  $\mathcal{R}$  se révèle ici inutile.

Nos définitions d'erreurs ont été inspirées de ces définitions de médiane et de moyenne. Le détail des calculs permet aussi de se convaincre que c'est pour la même raison que d'une part la moyenne minimise la fonction  $Q(x)$  et que d'autre part la droite qui minimise l'erreur quadratique passe par le point moyen du nuage.

Mais on doit observer aussi que cette propriété ne se transfère pas à la situation de l'erreur absolue. Dans le cas de notre exemple, le point médian a comme coordonnées  $(5; -4)$ , mais il n'est pas situé sur la droite passant par  $A$  et  $B$ , dont l'équation est

$$y = \frac{2}{13}x - \frac{68}{13}.$$

En réalité, c'est le fait que la droite minimisante passe toujours par deux parmi les trois points qui correspond le mieux à la notion de médiane telle qu'elle vient d'être décrite. Mais c'est une autre histoire...

Cela laisse bien deviner que nous n'avons soulevé qu'un tout petit pan du voile. Il y a beaucoup d'autres notions, résultats et problèmes qui s'articulent autour de cette question de la meilleure approximation d'un nuage de points par une droite. Bien sûr, la première chose à faire serait de généraliser tout ce qui précède à des nuages un

peu plus sérieux, comportant donc plus de trois points. Mais ça, le plus pénible, c'est de l'écrire : il n'y a sinon rien de neuf !

## Revenons-en au tabagisme

Tout ce que nous avons décrit ci-dessus s'applique évidemment aussi à notre problème initial de tabagisme. Le recours à EXCEL® permettra d'être bref !

Pour l'erreur absolue, on obtient la fonction

$$f(x) = 0,005x + 7$$

avec une erreur minimale égale à 8. On remarque que le Delaware, l'Indiana et le Montana sont situés dans l'alignement correspondant.

Pour l'erreur quadratique, on trouve la fonction

$$f(x) = 0,00542x + 6,908$$

avec une erreur minimale de 21,2881..., dont la racine carrée, plus significative ici, vaut 4,6139... Attention ! Ces beaux résultats restent à interpréter, et c'est ici que le mathématicien doit être encore plus prudent que devant une démonstration de 300 pages. En restant dans le seul cas des 6 couples de nombres de l'étude sur le tabagisme, la question n'est pas tant de savoir quelle droite est meilleure que l'autre, mais bien de comprendre pourquoi l'une ou l'autre mesure d'erreur est mieux adaptée au problème. Et pour cela, les techniques statistiques, mais aussi celles des sciences humaines, sont infiniment plus développées que ce petit article ne le laisse entrevoir. Pour qui voudrait en avoir une idée plus précise, on ne peut que vivement recommander la lecture des articles consacrés aux problèmes de santé publique dans le numéro spécial de la revue *la Recherche*, intitulé *Petits et Grands Nombres*, Hors-Série n° 13, Octobre-décembre 2003.

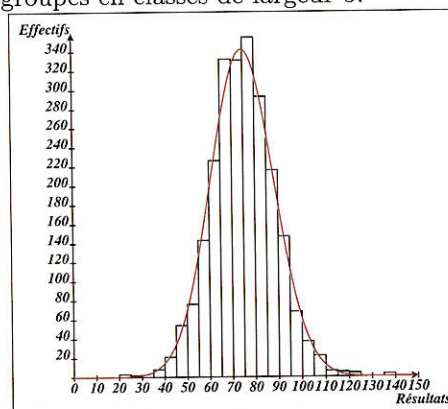
### Modéliser à l'aide d'une loi normale

Le tableau suivant résume les résultats obtenus en janvier 2005 par les 2352 élèves de 6<sup>e</sup> année ayant participé à l'éliminatoire de la maxi-olympiade Mathématique belge. Les scores  $y$  sont regroupés en classes de largeur 5.

De	à	Effectifs	De	à	Effectifs	De	à	Effectifs
20	24	3	60	64	226	100	104	37
25	29	2	65	69	332	105	109	22
30	34	0	70	74	331	110	114	6
35	39	8	75	79	355	115	119	5
40	44	21	80	84	293	120	124	4
45	49	54	85	89	216	125	129	0
50	54	76	90	94	147	130	134	0
55	59	143	95	99	68	135	139	3

Une loi normale est donnée par la fonction  $\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}}$  où  $m$  et  $\sigma$  sont respectivement la moyenne et l'écart-type des données considérées.

En superposant l'histogramme des données du tableau ( $m = 74,25$ ,  $\sigma = 13,72$ ) au graphique de la fonction normale correspondante (après normalisation des ordonnées), nous obtenons la figure ci-contre. L'approximation est suffisamment bonne pour admettre que les scores observés se répartissent effectivement suivant une loi normale.





# Calcul statistique d'une aire

Jean-Marc Desbonnez

## Prérequis

Pour réaliser le fichier décrit ci-dessous, il faut déjà avoir travaillé un peu avec EXCEL®. Les articles précédemment parus dans Maths-Jeunes : *Le déroulement de la spirale d'Archimède* (Math-Jeunes n°107) et *La loi du hasard* (Math-Jeunes n°108), détaillent tout ce qu'il faut savoir faire, et même plus. Des versions détaillées de ces articles sont disponibles sur le site de la SBPMef à l'adresse <http://www.sbpme.be>.

## Vous avez dit Monte-Carlo ?

La méthode de Monte-Carlo date du *projet Manhattan* qui, pendant la deuxième guerre mondiale de 1942 à 1945, a réuni les meilleurs physiciens et mathématiciens résidant aux Etats-Unis pour réaliser la première bombe atomique. Elle a été développée par Enrico Fermi, Nicholas Métropolis, John von Neumann et Stanislaw Ulam.

L'appellation de *Monte-Carlo* est tout à fait appropriée, puisque la méthode est fondée sur l'utilisation de séries de nombres aléatoires engendrées par ordinateur.

## La méthode de Monte-Carlo

La méthode de Monte-Carlo calcule très simplement une valeur approchée de l'aire d'une surface  $S$  délimitée par le graphique d'une fonction  $f(x)$  sur un intervalle  $[a; b]$ , l'axe des abscisses et les deux droites  $x = a$  et  $x = b$ . Ce procédé est indépendant des notions de dérivées et d'intégrales qui sont traditionnellement associées à ce genre de problème.

On suppose que la fonction  $f(x)$  est continue et bornée sur l'intervalle  $[a; b]$  et on note  $h$  un majorant de  $f(x)$  sur cet intervalle. Pour simplifier, on suppose encore que la fonction  $f(x)$  est positive.

Si on engendre un grand nombre de points dans le rectangle « de tir » déterminé par les quatre droites  $x = a$ ,  $x = b$ ,  $y = 0$  et  $y = h$ , et si ces points sont uniformément répartis dans ce rectangle, alors (cfr. Fig. 1) on doit avoir :

$$\frac{\text{Nombre de points en dessous de } y = f(x)}{\text{Nombre de points dans le rectangle}} \approx \frac{\text{Aire de } S}{\text{Aire du rectangle}}$$

En notant  $p$  le rapport  $\frac{\text{Nombre de points en dessous de } y=f(x)}{\text{Nombre de points dans le rectangle}}$ , on obtient ainsi une formule de calcul statistique de l'aire sous le graphique de  $f(x)$  :  $\text{Aire de } S \approx p \cdot \text{Aire du rectangle} = p \cdot (b - a) h$ .

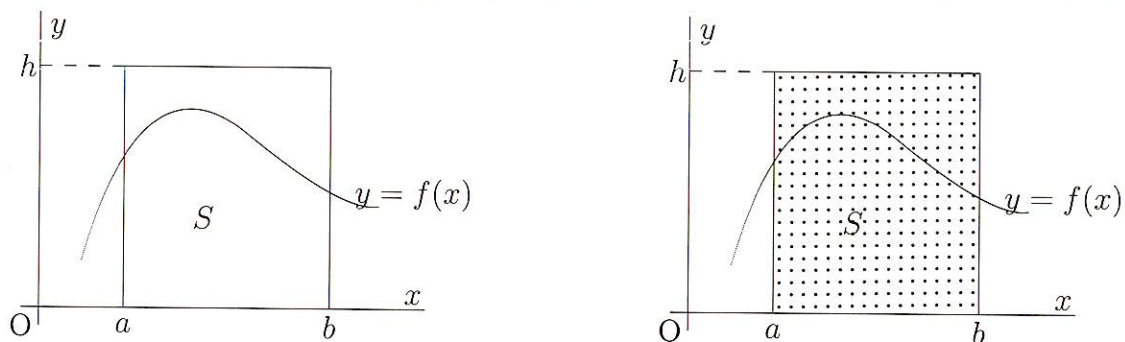


Figure 1 : Le graphique de la fonction, le rectangle « de tir » et une répartition (très) uniforme de points dans ce rectangle.



L'essentiel du travail consiste donc à évaluer le rapport  $p$  à partir des valeurs de  $f(x)$ . C'est évidemment ici que la simulation sous EXCEL<sup>®</sup> intervient. Pour lire la suite, nous vous conseillons vivement d'ouvrir un fichier EXCEL<sup>®</sup> et d'effectuer les manipulations décrites.

## La construction du fichier

L'objectif est de réaliser un fichier EXCEL<sup>®</sup> qui fournit une estimation d'une aire sous le graphique d'une fonction, mais qui soit très facile à adapter : à chaque changement de bornes des intervalles, nous voulons qu'il n'y ait qu'un minimum de changements à effectuer. Nous allons y parvenir en renommant certaines cellules, et en calculant les valeurs de la fonction à l'aide d'une table.

Pour fixer les idées, considérons la fonction  $f(x) = \sqrt{1-x^2}$  sur l'intervalle  $[-1; +1]$ , avec  $h = 1$ . Comme le graphique en question est celui d'un demi-cercle de rayon 1, nous savons d'avance que l'aire correspondante doit être égale à  $\frac{\pi}{2} = 1,57079\dots$ , ce qui doit nous permettre de tester la méthode.

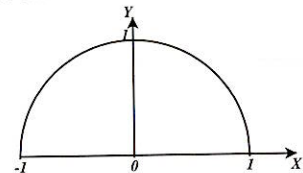


Figure 2

### Première étape : un nouveau nom pour certaines cellules

Ouvrons une feuille de calcul EXCEL<sup>®</sup>. Nous allons localiser dans la plage de cellules de B4 à B6 les informations relatives au « rectangle de tir » c'est-à-dire les valeurs de  $a$ ,  $b$  et  $h$ . Pour commencer, écrivons les commentaires **a:**, **b:** et **h:** en A4, A5 et A6.

Attribuons à présent des noms aux cellules B4 à B6. Pour attribuer un nom à la cellule B4, sélectionnons cette cellule et choisissons « Insertion/Nom/Définir » dans la barre de menus. Complétons le champ « Noms dans le classeur » par **a** et fermons la boîte de dialogue par OK. Désormais la cellule sélectionnée n'est plus repérée par l'adresse B4 mais bien par **a** : la variable  $a$  est devenue une *adresse de cellule* ! On renomme pareillement les cellules B5 et B6 respectivement en **b** et **h**.

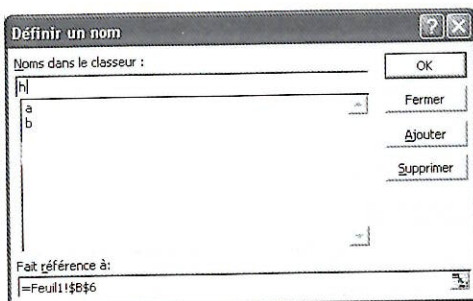


Figure 3

Toutes les cellules dont nous aurons besoin ont maintenant un nom suggestif, et de plus leur nouvelle adresse est devenue une référence *absolue* dans la feuille de calculs. Il reste à introduire les données : écrivons par exemple **-1** en B4, **1** en B5 et **1** en B6.

### Deuxième étape : construire le graphique de $f(x)$ et le rectangle de tir

Il est toujours intéressant de bien voir ce qu'il s'agit de calculer. Nous allons donc réaliser le graphique de la fonction  $f(x)$ , avec 100 points d'abscisses équidistantes sur l'intervalle  $[a; b]$ . Écrivons les commentaires **x** et **f(x)** en B9 et C9, et créons ensuite, de A11 à A111, la liste des nombres entiers de 0 à 100. De B11 à B111, introduisons la formule  $=a+A11*(b-a)/100$ .

Pour créer les valeurs correspondantes de la fonction  $f(x)$ , introduisons  $=\text{RACINE}(1-B11*B11)$  en C11 et recopions cette formule vers le bas jusqu'en C111.

	A	B	C
8			
9		x	f(x)
10			
11	0	-1	0
12	1	-0,98	0,19899749
13	2	-0,96	0,28
14	3	-0,94	0,34117444
15	4	-0,92	0,39191836
16	5	-0,9	0,43588989
17	6	-0,88	0,47497368
18	7	-0,86	0,51029403
19	8	-0,84	0,5425864
20	9	-0,82	0,57236352
21	10	-0,8	0,6
22	11	-0,78	0,62577951
23	12	-0,76	0,64992307
24	13	-0,74	0,67260687

Figure 4



Un modèle du fichier décrit dans cet article est disponible sur le site de la SBPMef à l'adresse <http://www.sbpme.fr>, sous le nom de MonteCarloMJmanuel.xls. Un autre fichier réalisant le même calcul mais basé sur l'emploi de tables de valeurs est également disponible sous le nom de MonteCarloMJtables.xls.

abscisse	ordonnée
=a	0
=b	0
=b	=h
=a	=h
=D10	=E10

Figure 5 : Les coordonnées du rectangle de tir.

## Pour effectuer un collage spécial

- Sélectionner les données à importer, ici la plage de cellules D10 à E14 ;
- choisir « Edition/Copier » ;
- sélectionner la fenêtre graphique dans laquelle on veut importer ces données ;
- choisir « Edition/Collage Spécial » ;
- vérifier que la boîte de dialogue qui s'ouvre alors est complétée comme dans la figure 6 ci-dessous.

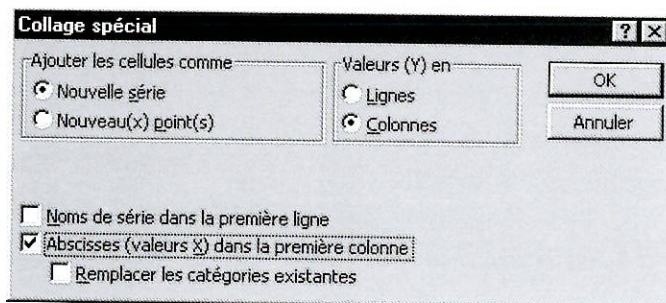


Figure 6 : La boîte de dialogue d'un collage spécial.

Créons maintenant les coordonnées des sommets du rectangle de tir dans la plage de cellules D10 à E14. Il suffit de calculer ces coordonnées en fonction de  $a$ ,  $b$  et  $h$ , ce qui n'offre pas de difficultés. (Figure 5.)

Les coordonnées du premier sommet sont reproduites en cinquième position pour pouvoir fermer (graphiquement) le rectangle.

Il reste enfin à vraiment *visualiser* tout cela. Le graphique de  $f(x)$  est réalisé en suivant la procédure standard, à partir de « Insertion/Graphique... » dans la barre de menus, et en choisissant le graphique en « Nuages de Points ». Le rectangle y est alors superposé par un *collage spécial*. La façon de réaliser un collage spécial est rappelée en marge.

Finalement, il faut encore modifier le type de graphique pour relier entre eux ces 5 points, ce qui se réalise en sélectionnant d'abord l'un de ceux-ci (toute la série suit) et, après un clic droit, en sélectionnant « Format de séries de données » dans le menu déroulant, puis dans l'onglet Motifs, « Trait automatique ». La partie gauche de la figure 7 plus loin montre ce qui est déjà présent à l'écran à la fin de ces deux premières étapes.

## Troisième étape : la génération des points dans le rectangle et le test de Monte-Carlo

Nous voulons répartir des points uniformément dans le rectangle de tir ; une quantité de 1000 points est suffisante pour un ordre de grandeur raisonnable dans l'approximation, ainsi que dans le temps de calcul.

Chacun de ces 1000 points est défini par son abscisse  $v$  et son ordonnée  $w$ , qui doivent être des nombres pris au hasard, respectivement dans les intervalles  $[a; b]$  et  $[0; h]$ . Ce que nous appelons le *test de Monte-Carlo* revient à comptabiliser les points vérifiant la condition  $f(v) < w$ .

On introduira dans la plage K11 à K1010 des nombres  $v$ , pris au hasard entre  $a$  et  $b$ . Les valeurs  $f(v)$  seront calculées dans la plage L11 à L1010 et d'autres nombres  $w$ , pris au hasard entre 0 et  $h$  dans la plage M11 à M1010. Il suffira en fait d'écrire  $=a+(b-a)*ALEA()$  en K11,  $=RACINE(1-K11*K11)$  en L11,  $=h*ALEA()$  en M11, puis de recopier ces trois éléments vers le bas jusqu'à la ligne 1010.

Désormais, tout est clair : il reste à écrire en N11 et à reproduire jusqu'en N1010 la formule  $=SI(M11<L11;1;0)$  qui affiche « 1 » si la condition  $f(v) < w$  est vérifiée, et « 0 » sinon.

Pour réaliser le calcul d'aire proprement dit, il ne reste plus qu'à compter ! On écrira, par exemple en I18 :  $=SOMME(N11:N1010)$  et en I20 :  $=I18*(b-a)*h/1000$ .



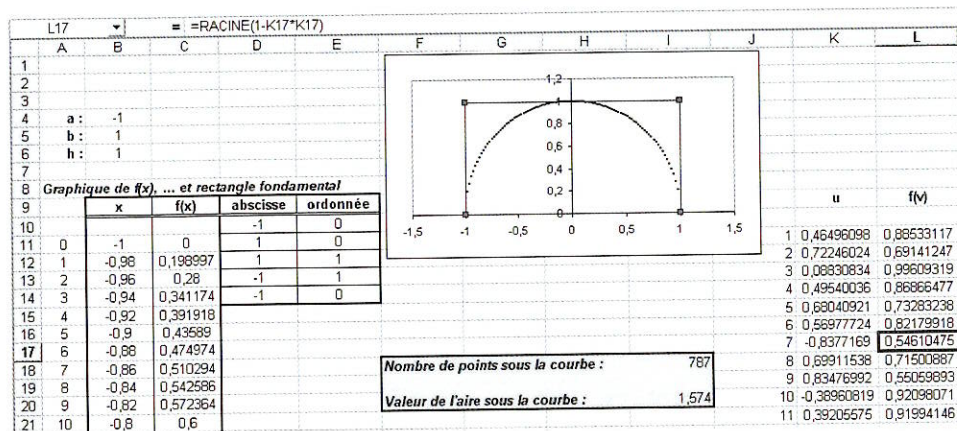


Figure 7 : Une partie du fichier à l'écran

L'exemple que livre la figure 7 suggère ainsi que  $\frac{\pi}{2} \approx 1,574$  ce qui n'est pas trop mauvais comme approximation ...

## Epilogue (hasardeux...)

Quand un fichier est fini, la première chose à faire est de jouer un peu avec la bête, et de tester l'estimation qu'elle propose dans des cas où la réponse est connue. N'importe quel professeur de mathématiques peut certainement fournir toute une bibliothèque de cas de ce genre.

Ces premières expériences montrent que le fichier ne donne qu'une approximation assez fluctuante, et donc relativement grossière de l'aire à évaluer. Mais il le fait très vite (!), y compris pour de « vilaines » fonctions. Par exemple, et pour ceux qui savent déjà ce qu'est une intégrale, ce fichier permet d'estimer des intégrales pour des fonctions dont les primitives ne se calculent pas en termes de fonctions élémentaires et cela, même si les fonctions à intégrer ont des comportements assez sauvages. Un exemple parmi d'autres est celui du calcul de  $\int_0^1 \left| \sin \frac{1}{x} \right| dx = \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^1 \left| \sin \frac{1}{x} \right| dx$ .

Une question qui se pose est celle de la précision de l'évaluation. On démontre dans les cours de probabilités que pour augmenter cette précision d'un facteur  $\frac{1}{N}$  dans une méthode de Monte-

Carlo, il faut multiplier le nombre d'expériences d'un facteur  $N^2$ ; en clair, pour gagner une décimale, il faut à chaque fois multiplier le nombre d'expériences par 100. Chaque expérience correspondait à une valeur de  $v$ ,  $f(v)$  et  $w$ , c'est le nombre de lignes utilisées dans le tableur qui est multiplié par 100. Voilà qui nécessite un tableur particulièrement rapide et puissant! À moins d'utiliser des macros visual basic qui permettent une approche toute différente... mais qui débordent du cadre de cet article.

Pour finir, une petite question toute simple. Que se passe-t-il si on fait tourner le fichier *en oubliant* certaines hypothèses de départ, par exemple si on choisit d'étudier l'aire sous le graphique de la fonction  $f(x) = \frac{1}{x}$  avec  $a = 0$  et  $b = 2$ ? Cette fonction n'est pas bornée sur l'intervalle  $[0; 2]$ , mais quelques essais montrent que n'importe quel choix d'une valeur de  $h$  donne un résultat qui semble relativement stable. De plus, avec  $h = 100$ , on trouve par exemple après  $5 \cdot 10^6$  itérations, une aire égale à 6,2998. Est-ce dû au hasard, ou y a-t-il une bonne raison?

### Un petit conseil

Il y a moyen d'empêcher EXCEL® de mettre à jour tout le classeur à chaque modification d'une cellule. Il suffit de choisir « Outils/Options... » dans la barre de menus, et dans l'onglet « Calcul/Mode de calcul » de cocher l'option « Sur ordre ». Les nombreux calculs ne seront alors mis à jour que lorsqu'on pressera la touche F9 tout en haut du clavier.

Connaissez-vous

## L'écho des savants

Ce bimestriel est publié par les Jeunesses scientifiques de Belgique. Pour plus d'informations, consultez le site

www.jsb.be



# Histoire et mathématiques des tables de mortalité

Daniel Justens

William Petty  
1623 – 1687

Proportion de décès par  
groupe d'âge ( $Q_x$ ), (période  
1996–1998)

	Total	Hommes	Femmes
0	0,003854	0,004327	0,003361
1	0,000960	0,001058	0,000857
2	0,000343	0,000359	0,000430
3	0,000296	0,000374	0,000214
4	0,000214	0,000218	0,000210
5	0,000188	0,000251	0,000122
6	0,000198	0,000202	0,000193
7	0,000188	0,000199	0,000175
8	0,000197	0,000225	0,000167
9	0,000126	0,000155	0,000097
10	0,000109	0,000137	0,000080
11	0,000157	0,000154	0,000161
12	0,000088	0,000078	0,000099
13	0,000170	0,000128	0,000216
14	0,000174	0,000258	0,000085
15	0,000274	0,000387	0,000154
16	0,000409	0,000532	0,000279
17	0,000564	0,000741	0,000378
18	0,000574	0,000813	0,000324
19	0,000559	0,000746	0,000364
20	0,000723	0,001065	0,000368

Source : Institut  
National de Statistique

## Notes

(<sup>1</sup>) Les critères de regroupement sont l'âge et le sexe (depuis 1993) mais on constate aussi des différences importantes de longévité entre les personnes qui s'assurent et celles qui ne s'assurent pas ce qui est plus étonnant. De plus, le législateur impose de tarifier séparément les contrats en cas de vie et les contrats en cas de décès.

(<sup>2</sup>) Comme la surmortalité des jeunes de sexe masculin entre 15 et 25 ans.

## Qu'est-ce qu'un contrat d'assurance-vie ?

Un contrat d'assurance vie est un accord entre un particulier et un organisme assureur garantissant à un bénéficiaire le paiement par l'organisme de capitaux convenus à la signature. Le paiement se fait soit en cas de survie du particulier à un ou plusieurs moments dans le futur (capital différé ou rente), soit en cas de décès de celui-ci. En contrepartie, le client s'engage au paiement de primes.

Pour tarifier ce type de contrat, la compagnie d'assurances doit construire un modèle rationnel, ajusté aux observations, des probabilités de décès pour chaque sous-ensemble homogène(<sup>1</sup>) de personnes couvertes. La compagnie doit également développer des modèles financiers de manière à pallier les effets de l'érosion monétaire. Voyons en quoi consistent les modèles viagers.

## Les observations

Comment observer le *risque de décès* d'un individu ? Première constatation : l'âge du décès est une variable continue. Pour la traiter, il faut procéder à des regroupements par *classes* et on choisit naturellement de travailler par année. Mais les choses sont moins évidentes qu'il n'y paraît. Pour une seule observation de la proportion de décès des hommes ou des femmes dans leur 50<sup>e</sup> année (entre 49 et 50 ans), il faut deux ans : on observe chaque individu entre deux dates anniversaires successives pour assurer l'homogénéité des observations et les anniversaires se distribuent presque uniformément sur l'année civile. De plus, les taux de décès évoluent rapidement au cours du temps, nos populations ayant pris l'agréable habitude de vivre de plus en plus longtemps, ce qui rend toute observation trop ancienne caduque. Il faut donc se contenter de peu de données pour l'ajustement.

Le lecteur intéressé par des données expérimentales peut se connecter au site de l'Institut National de Statistique : [http://statbel.fgov.be/figures/download\\_fr.asp](http://statbel.fgov.be/figures/download_fr.asp).

Mais les fréquences observées de décès ne sont pas directement utilisées par les actuaires. Pour pallier certaines variations surprenantes(<sup>2</sup>), il convient de construire un modèle probabiliste rationnel ajusté aux observations. Voyons comment, historiquement, sont apparues les premières tables théoriques de mortalité.



## De premiers recensements

La première civilisation à avoir développé une structure étatique suffisante pour gérer un cadastre est celle de l'Égypte. Dès la II<sup>e</sup> dynastie (2925 – 2700 BC), on trouve des traces écrites d'un recensement. Sous la IV<sup>e</sup> dynastie (2625 – 2510 BC), la biographie de Meten<sup>(3)</sup> nous apprend que ce dernier a pu acheter une rente de deux cents pains par jour. Peut-on pour autant en conclure dès cette époque à l'existence d'une mathématique viagère<sup>(4)</sup> ? En fait, ce contrat est exceptionnel. On n'en retrouvera aucun autre exemple dans l'Antiquité et il faudra attendre 4000 ans pour voir apparaître les premières études démographiques.



John Graunt  
1620–1674

## Impensable croissance infinie

Ce qui manque, c'est la notion de progression géométrique<sup>(5)</sup>. Dans l'univers aristotélécien où la sphère des étoiles « fixes » limite l'espace à quelques milliers de rayons terrestres, une croissance indéfinie de ce type ne peut être envisagée. Paradoxalement, c'est William Petty, auteur probable de la première table de mortalité moderne<sup>(6)</sup>, qui va également justifier l'impossibilité des progressions. Se référant aux textes sacrés, il note dans son *Autre Essai d'Arithmétique politique* (1682), se livrant à d'audacieuses spéculations sur les temps de doublement de la population :

*Si le nombre d'acres que mesure la partie habitable de la terre est inférieur à 50 milliards, si une population de 20 milliards est trop considérable pour pouvoir être nourrie par ce nombre d'acres [...], alors en six redoublements de la population (ce qui se produira dans 2 000 ans), les 320 000 000 habitants dépasseront le chiffre en question de 20 milliards<sup>(7)</sup>. Et alors, suivant la prédiction des écritures, il y aura des guerres et de grands massacres.*

La progression géométrique ne pouvait exister puisqu'elle conduisait inévitablement à l'apocalypse. Malgré tout, on soupçonne fort Petty d'avoir œuvré avec le londonien, John Graunt à la publication en 1662 des *Natural and Political Observations made upon the Bills of Mortality*, fondant ainsi l'analyse démographique<sup>(8)</sup> :

*Puisque nous avons trouvé que sur 100 conceptions prises au départ, à peu près 36 n'atteignent pas l'âge de 6 ans, et que peut-être une seule survit à 76 ans, ayant sept décennies entre 6 et 76 ans, nous avons recherché six moyennes proportionnelles entre 64, ceux qui sont encore vivants à 6 ans, et l'unique survivant à 76 ans. [...] De là, il s'ensuit que sur 100 personnes conçues, il en reste,*

### Note

<sup>(3)</sup> Jacques PIRENNE, *Histoire de la civilisation de l'Égypte ancienne*, tome 1 p. 176, Baco-nnière-Renaissance du Livre, Neuchâtel-Bruxelles, 1963.

<sup>(4)</sup> Pour apprendre l'arithmétique égyptienne : voir GILLAIN, *La science égyptienne : l'arithmétique au Moyen Empire*, Edition de la Fondation égyptologique Reine Élisabeth, Bruxelles, 1927.

<sup>(5)</sup> Ou, en continu, l'exponentielle.

<sup>(6)</sup> Cette dernière est très souvent attribuée à John Graunt.

<sup>(7)</sup> Autrement dit le produit de 320 millions par  $2^6$  est supérieur à 20 milliards.

<sup>(8)</sup> Voir l'ouvrage d'Hervé Le Bras : *Naissance de la mortalité*, Gallimard, Le Seuil, 2000.





Edmund Halley  
1656–1742

au bout de six années pleines	64
au bout de 16 ans	40
au bout de 26 ans	25
au bout de 36 ans	16
au bout de 46 ans	10
au bout de 56 ans	6
au bout de 66 ans	3
au bout de 76 à 80 ans	0

On a beaucoup spéculé sur l'origine de ces chiffres. Plusieurs méthodes de reconstitution ont été proposées, basées sur des séries géométriques de raison 0,63, ou 0,625<sup>(9)</sup>, mais aucune ne donnait une reconstitution correcte.

Hervé Le Bras nous livre sa solution<sup>(10)</sup>. Le chiffre initial de 64 met sur la piste : il représente six multiplications successives par 2. Or la *duplicatio* (multiplication par 2) et la *manducatio*, (division par 2) étaient considérées au XVII<sup>e</sup> siècle comme des opérations au même titre que nos quatre opérations usuelles. Pour prendre 64 % d'un nombre donné, il suffit de le multiplier six fois de suite par 2 et de le diviser par 100, en supprimant les deux derniers chiffres. Si l'on opère ainsi, et si l'on arrondit par suppression de la partie décimale comme on le faisait alors, on reconstitue *presque* la table des *Observations*.

La technique repose sur des arrondis et des multiplications simples ; ces opérations s'effectuent mentalement, la notation décimale n'étant pas encore systématique. La première table de mortalité fut donc une série géométrique décroissante calculée de manière approximative.

La voie était tracée et, au cours des XVII<sup>e</sup> et XVIII<sup>e</sup> siècles, parurent de multiples tables ; citons entre autres celles du célèbre Astronome anglais Edmund Halley, l'homme de la comète, qui proposa en 1693 une table de mortalité de la ville de Breslau (actuellement Wrocław, en Pologne). Mais il fallut attendre les travaux de Benjamin Gompertz pour voir naître le premier modèle.

## Enfin des mathématiques !

Graunt et Petty ont fait œuvre de précurseurs en décrivant, sur base d'observations locales, l'évolution d'une population fictive. Mathématisons la démarche et notons  $\ell_0$  l'effectif d'une « population » initiale<sup>(11)</sup> (Pour Graunt,  $\ell_0 = 100$  mais aujourd'hui, on prend  $\ell_0 = 10^5$  ou  $10^6$ ). En notant  $p_x$  la probabilité<sup>(12)</sup> pour un individu d'âge  $x$  d'être encore en vie un an plus tard<sup>(13)</sup>, on construit successivement les nombres théoriques de « survivants ».

$$\text{Pour } x \in \mathbb{N}, \quad \ell_{x+1} = p_x \cdot \ell_x. \quad (A)$$

C'est cet ensemble de valeurs  $\ell_x$  que l'on appelle « table de mortalité ». Abandonnons le point de vue discret et considérons la population  $\ell_x$  comme une fonction continue et dérivable<sup>(14)</sup> de l'âge  $x$ .

### Notes

<sup>(9)</sup> 5/8 : cette dernière fut avancée par Karl Pearson, l'un des fondateurs de la statistique mathématique.

<sup>(10)</sup> L'invention des concepts en démographie, *Dossier Pour la Science : Les mathématiques sociales*, 1999.

<sup>(11)</sup> Tout se passe en théorie comme si  $\ell_0$  personnes naissent en même temps et que l'on observe l'évolution de l'effectif de cette *population* au fil du temps.

<sup>(12)</sup> Nous verrons plus loin comment ajuster un modèle probabiliste aux observations.

<sup>(13)</sup> On note aussi  $q_x = 1 - p_x$  la probabilité de décès entre l'âge  $x$  et l'âge  $x + 1$ .

<sup>(14)</sup> La dérivation est étudiée normalement en cinquième année. Le lecteur ayant des difficultés à ce sujet peut « avancer » à la formule (B) qui exprime le résultat des calculs.



Calculons le nombre de décès par unité de temps entre les âges  $x$  et  $x + \Delta x$  :

$$\frac{\ell_x - \ell_{x+\Delta x}}{\Delta x}$$

On appelle *taux instantané de décès*<sup>(15)</sup> à l'âge  $x$ , et on note  $\mu_x$  le taux de décès par unité de temps au voisinage de l'âge  $x$ . La fonction  $\ell_x$  peut être calculée à partir de  $\mu_x$ . En effet, on peut écrire<sup>(16)</sup> :

$$\mu_x = \lim_{\Delta x \rightarrow 0} \frac{1}{\ell_x} \cdot \frac{\ell_x - \ell_{x+\Delta x}}{\Delta x} = -\frac{d}{dx} [\ln \ell_x]$$

Les tables de mortalité actuelles se fondent sur une modélisation interprétée de ce taux instantané. Benjamin Gompertz constata (en 1825), outre une énorme mortalité infantile suivie d'une décroissance lors de l'adolescence, que les taux de décès des adultes croissaient exponentiellement. Il posa donc ( $20 \leq x \leq 80$ ) :

$$\mu_x = \alpha c^x$$

En 1860, William Makeham observa que l'adjonction d'une constante additive améliorerait considérablement l'adéquation du modèle aux observations. Il proposa une représentation basée sur l'hypothèse d'un risque accidentel constant quel que soit l'âge et d'un risque exponentiel lié au vieillissement :

$$\mu_x = A + \alpha c^x$$

$A$  représente le risque accidentel ( $A > 0$ ),  $\alpha$  le risque initial lié à la population considérée ( $\alpha > 0$ ) et  $c$  le coefficient d'aggravation du taux de décès par année ( $c > 1$ ).

Ce modèle permet facilement de construire une table  $\ell_x$ . Si

$$\frac{d}{dx} [\ln \ell_x] = -A - \alpha c^x,$$

en intégrant, on obtient successivement :

$$\ln \ell_x = -Ax - \frac{\alpha}{\ln c} c^x + \ln k$$

$$\ell_x = k \cdot e^{-Ax} \cdot e^{-\frac{\alpha}{\ln c} c^x} \quad (\text{B})$$

La constante d'intégration  $k$  a dans ce contexte une interprétation : elle exprime le choix arbitraire de la population fictive initiale  $\ell_0$  (attention : les deux quantités sont liées mais différentes). Depuis, de nombreux modèles sont venus enrichir la théorie mais le législateur belge impose toujours la modélisation de Makeham aux assureurs.

La paramétrisation du modèle n'est pas trop délicate : on peut calculer théoriquement les  $p_x$  à partir de (A) et (B), un double passage aux logarithmes permet de retrouver du linéaire et l'ajustement aux observations des fréquences observées de survie se fait par moindres carrés. Dans le *Moniteur Belge* du 31 décembre 1992 (page 27 887), le législateur belge a fixé le modèle de tarification en partant de la paramétrisation qui suit, basée sur le seul clivage du sexe, en autorisant



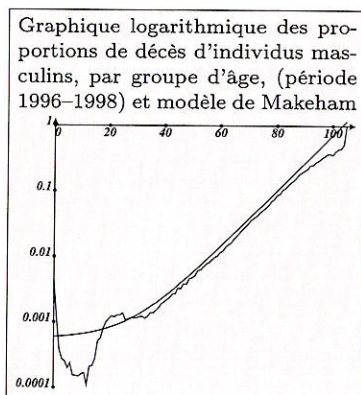
Benjamin Gompertz  
1779 – 1865

## Notes

<sup>(15)</sup> Le *taux de décès* entre l'âge  $x$  et l'âge  $x + \Delta x$  est la proportion de décès :  $\frac{\ell_x - \ell_{x+\Delta x}}{\ell_x}$ . Le *taux de décès par unité de temps* entre  $x$  et  $x + \Delta x$  s'obtient en divisant l'expression précédente par  $\Delta x$ . Le *taux instantané de décès par unité de temps* au voisinage de  $x$  est la limite du taux de décès par unité de temps lorsque  $\Delta x$  tend vers 0.

<sup>(16)</sup>  $\ln$  est le logarithme népérien (voir page 27).





## Notes

(<sup>17</sup>) Cette comparaison a un sens car la probabilité de décès pour un individu d'âge  $x$  pendant un intervalle de durée  $\Delta t$  est en première approximation égale à  $\mu_x \Delta t$ .

(<sup>18</sup>) On observera la forte décroissance de la mortalité infantile et la « bosse » des 15 – 25 ans due aux accidents d'auto et aux suicides...

(<sup>19</sup>) Il s'agit évidemment d'une majoration...

(<sup>20</sup>) Chaque assureur couvre plusieurs centaines de milliers de personnes.

une marge de sécurité fonction du type de couverture envisagée (vie ou décès) :

	$c$	$\alpha$	$A$
Hommes	1.103798111	0.000037074	0.000558453
Femmes	1.119312877	0.000007255	0.000330325

Dans quelle mesure les tables construites sur base de ces valeurs sont-elles bien ajustées à la réalité observée? Afin de pouvoir en juger, le graphique ci-contre représente à la fois le modèle type *Makeham* imposé pour les hommes (fonction  $\mu_x$ ) et les observations empiriques des probabilités de décès(<sup>17</sup>) ( $q_x$ ) relatives à la période 1996–1998(<sup>18</sup>).

La table  $\ell_x$  construite, le calcul est débarrassé de son contexte probabiliste et devient algébrique. La prime *pure* est transformée(<sup>19</sup>) en prime *commerciale* par prise en compte des frais annexes et en assurant une marge de sécurité suffisante pour l'assureur.

## Conclusions

On l'a constaté, la modélisation des tables de mortalité laisse peu de place au calcul des probabilités et c'est aussi le cas de la mathématique de l'assurance vie. La paramétrisation de la fonction  $\mu_x$ , ajustée aux observations par une simple régression, fait disparaître en apparence l'aléatoire du problème. La table de mortalité, une fois construite, est considérée comme un référent statique. Les lois des grands nombres(<sup>20</sup>) autorisent ce travail de tarification quasiment déterministe et garantissent (au moins à court terme) la rentabilité du secteur.

## Solutions des jeux

## 1. Sudoku

## A. Grille facile

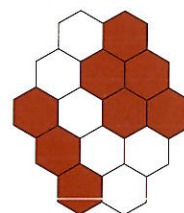
5	3	4	6	7	8	9	1	2
6	7	2	1	9	5	3	4	8
1	9	8	3	4	2	5	6	7
8	5	9	7	6	1	4	2	3
4	2	6	8	5	3	7	9	1
7	1	3	9	2	4	8	5	6
9	6	1	5	3	7	2	8	4
2	8	7	4	1	9	6	3	5
3	4	5	2	8	6	1	7	9

## B. Grille difficile

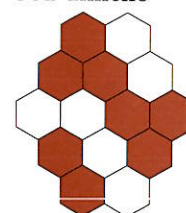
5	3	8	9	4	2	1	7	6
9	2	7	1	8	6	5	4	3
1	6	4	5	7	3	8	9	2
2	7	1	6	5	9	4	3	8
3	4	9	8	2	7	6	1	5
8	5	6	4	3	1	7	2	9
7	8	5	3	9	4	2	6	1
4	1	3	2	6	8	9	5	7
6	9	2	7	1	5	3	8	4

## 2. Les hexagones

## Jeu facile



## Jeu difficile



## 3. Pyramide magique

30	8	9	27	26
29	23	16	21	11
12	18	20	22	28
15	19	24	17	25
1	32	31	13	10

## 4. Les produits croisés

A.	9	1	7
	5	3	2
	4	6	8

B.	$x$	4	$x - 1$
	$2x + 1$	$x$	$3x$
	$2x - 1$	2	$x + 1$

## 5. Anneaux olympiques

Aux symétries près, il existe quatre solutions :

A	B	C	D	E	F	G	H	I
8	3	7	1	6	4	5	2	9
7	6	5	2	3	8	1	4	9
7	6	5	2	8	3	1	9	4
5	9	2	3	4	7	1	6	8

## 6. Anneaux olympiques et fantômes

Appariements :

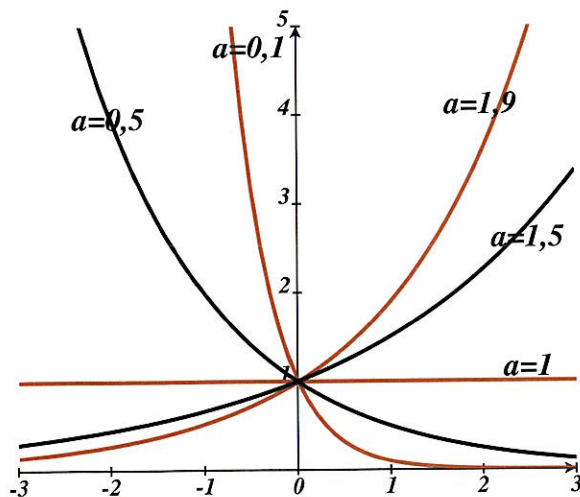
A-5, B-2, C-1, D-4, E-3.



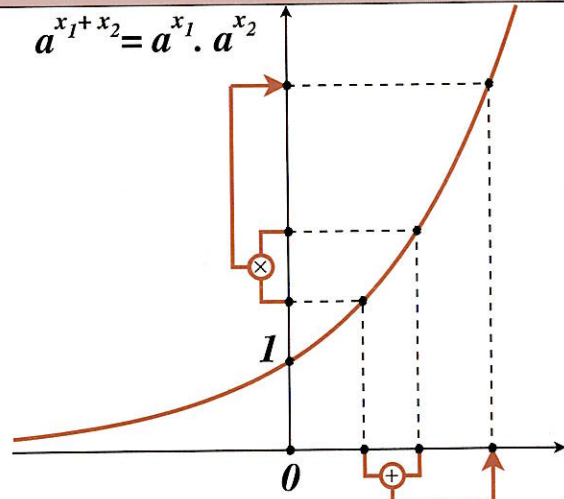
## LES fonctions exponentielles

Si  $a$  est un réel strictement positif, la notation  $a^x$  a un sens et désigne un réel strictement positif quel que soit le réel  $x$ .

La fonction  $x \mapsto a^x$  est la fonction exponentielle de base  $a$ . On note souvent  $\exp_a(x) = a^x$ .



Croissance et décroissance

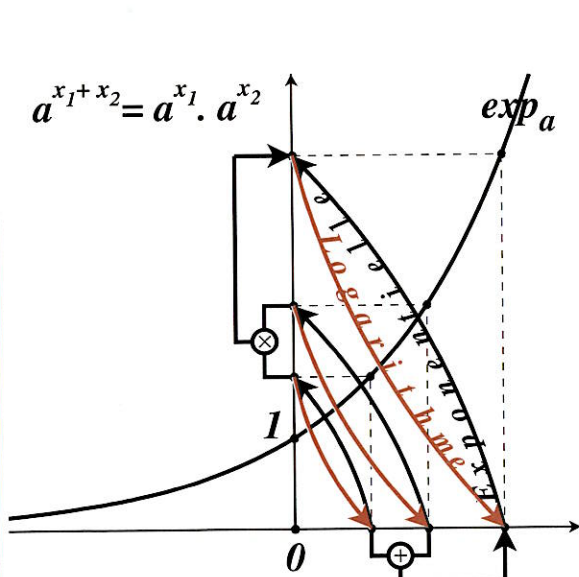


L'addition est transformée en multiplication

## LES fonctions logarithmes

Les fonctions logarithmes  $\log_a$  sont les réciproques des fonctions exponentielles  $\exp_a$ .

Le graphique de la fonction  $\log_a$  s'obtient en retournant le graphique de la fonction  $\exp_a$  autour de la 1<sup>re</sup> bissectrice. Puisque  $a^0 = 1$  et  $a^1 = a$ , on a  $\log_a(1) = 0$  et  $\log_a(a) = 1$ .



$$\log_a(y_1 \cdot y_2) = \log_a(y_1) + \log_a(y_2)$$

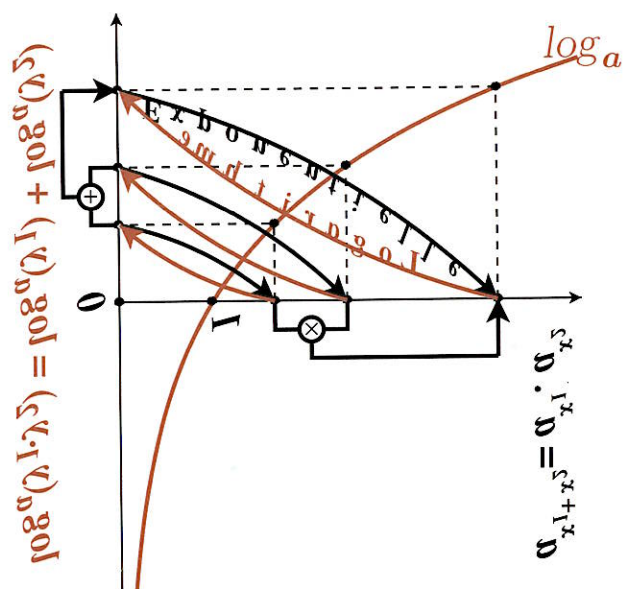
$$\forall x, \forall y > 0 :$$

$$y = \exp_a(x) \Leftrightarrow x = \log_a(y),$$

$$\log_a(a^x) = x,$$

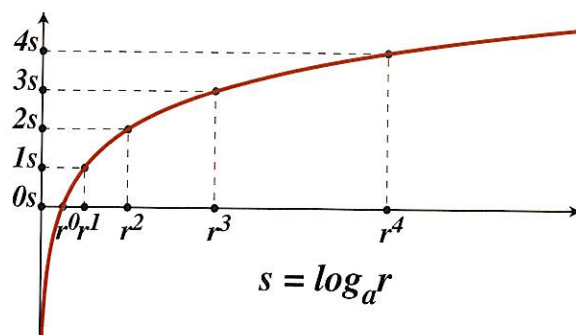
$$a^{\log_a(y)} = y$$

La multiplication est transformée en addition.



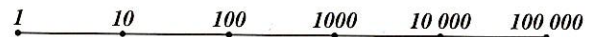


## Les échelles logarithmiques

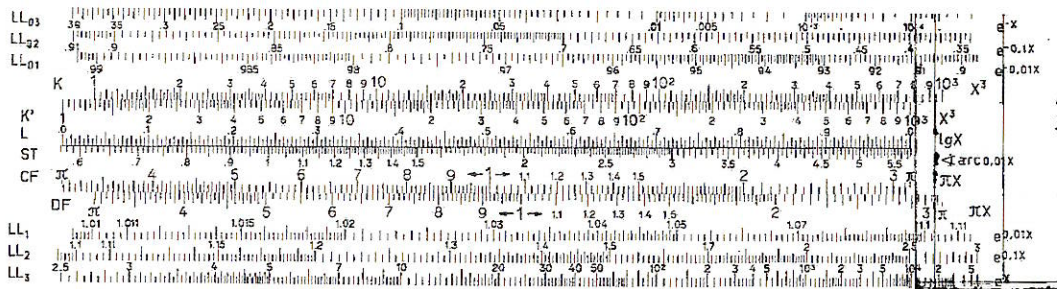


Toute fonction logarithme transforme une suite géométrique en une suite arithmétique.

Une échelle logarithmique sur une droite s'obtient en graduant cette droite à l'aide de nombres en progression géométrique. Les marques de graduation ont pour abscisses les logarithmes des termes de la progression.



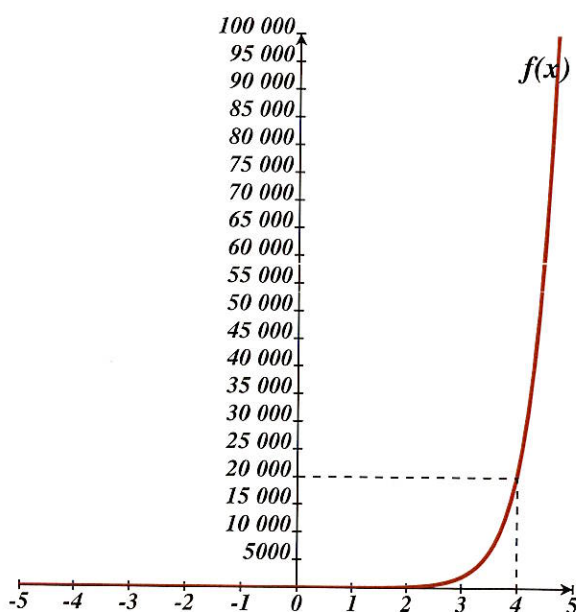
Une échelle logarithmique basée sur une progression géométrique de raison 10.



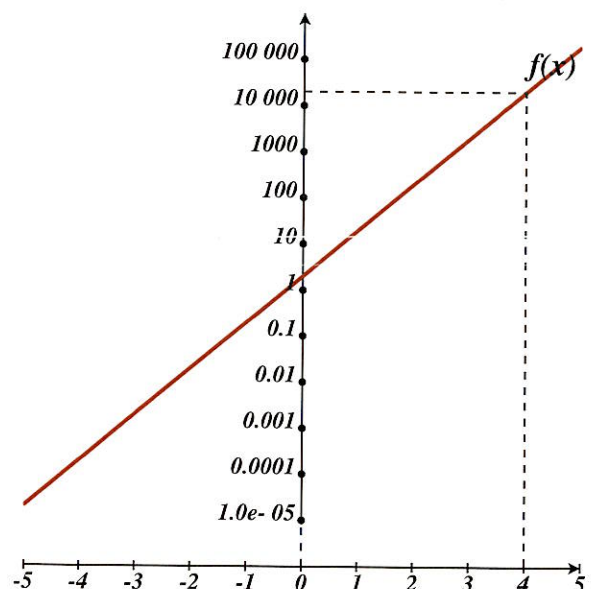
Les règles à calcul utilisées avant l'invention des calculatrices électroniques étaient graduées à l'aide d'échelles logarithmiques.

## Les graphiques logarithmiques

Dans un graphique logarithmique, l'axe des  $y$  est muni d'une échelle logarithmique. Un tel graphique est utile pour représenter des fonctions dont la variation est très importante. Le graphique logarithmique d'une fonction exponentielle est une droite. La figure de droite ci-dessous est un graphique logarithmique représentant la même fonction  $f(x) = 2 \times 10^x$  que le graphique cartésien de gauche.



Le graphique usuel ne fournit d'information numérique utilisable que pour  $x$  entre 2 et 4.



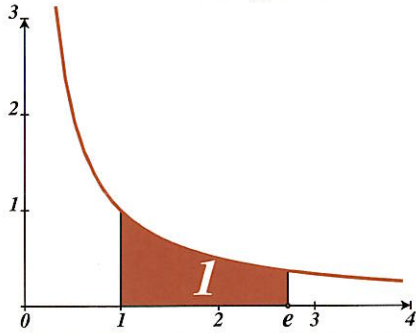
Le graphique fournit de meilleures approximations numériques.



## Le nombre $e$

Le nombre  $e$  défini par C. Huygens<sup>(21)</sup> :

$e$  est le nombre réel tel que l'aire comprise entre l'axe des  $x$ , l'hyperbole d'équation  $y = \frac{1}{x}$  et les verticales d'abscisse 1 et  $e$  est égale à 1.



D'autres définitions équivalentes de  $e$  :

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots$$

$$e = 2 + \frac{1}{1 + \frac{1}{2 + \frac{1}{3 + \frac{1}{4 + \frac{1}{5 + \frac{1}{6 + \dots}}}}}}$$

et enfin, une valeur approchée de  $e$  (due à Euler<sup>(21)</sup>) :

$$e \approx 2,718\,281\,828\,459\,045\,235$$

Les factorielles :

$$0! = 1, \quad 1! = 1, \quad 2! = 1 \times 2, \quad 3! = 1 \times 2 \times 3 = 6, \quad \dots, \quad n! = 1 \times 2 \times 3 \times \dots \times n$$

et leur relation avec les nombres  $e$  et  $\pi$  :

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

L'erreur relative commise en remplaçant  $15!$  par  $\sqrt{2 \times 15 \times \pi} \left(\frac{15}{e}\right)^{15}$  est inférieure à 1%.

## LA fonction exponentielle et LA fonction logarithme népérien<sup>(21)</sup>.

LA fonction exponentielle **exp** est la fonction exponentielle  $\exp_a$  pour laquelle  $a = e$ .

LA fonction logarithme népérien **ln** est la fonction réciproque de LA fonction exponentielle **exp**.

Comme toute fonction logarithme,  $\ln(x)$  n'est défini que pour  $x > 0$ .

$$\ln(e^x) = x \quad e^{\ln(x)} = x$$

Pour le lecteur qui a déjà rencontré le concept de « dérivée d'une fonction » :

Notons  $k_a$  la valeur de la dérivée en 0 de la fonction  $\exp_a$ . Alors la dérivée de cette fonction en une valeur quelconque  $x_0$  de  $x$  vaut

$$\lim_{\Delta x \rightarrow 0} \frac{a^{x_0 + \Delta x} - a^{x_0}}{\Delta x} = \lim_{\Delta x \rightarrow 0} a^{x_0} \cdot \frac{a^{\Delta x} - a^0}{\Delta x} = k_a \cdot a^{x_0}$$

La fonction dérivée de  $\exp_a$  est donc  $k_a \cdot \exp_a$ . On peut montrer que  $k_a = \ln a$ . Donc

$$(a^x)' = \ln(a) \cdot a^x$$

En particulier, puisque  $\ln(e) = 1$ , la fonction  $\exp(x) = e^x$  est sa propre dérivée. En dérivant les deux membres de l'expression  $e^{\ln(x)} = x$ , on obtient  $e^{\ln(x)} \cdot (\ln'(x)) = 1$  donc

$$\ln'(x) = \frac{1}{x}$$

<sup>(21)</sup> Christiaan Huygens, 1629–1695 ; Léonard Euler, 1707–1783 ; John Neper, ou Napier, 1550–1617.



# Jeux

Yolande Noël-Roch

## 1. Sudoku

**Les règles du jeu.** Un « grand carré »  $9 \times 9$  est partagé en neuf « petits carrés »  $3 \times 3$ . Quelques chiffres sont donnés et le jeu consiste à compléter les cases de manière à ce que les chiffres de 1 à 9 apparaissent chacun une et une seule fois dans

- chaque ligne du grand carré,
- chaque colonne du grand carré,
- chacun des neuf petits carrés.

Voici un exemple : partant de

6	1	4	5		
	8	3	5	6	
2					1
8		4	7		6
	6			3	
7		9	1		4
5					2
	7	2	6	9	
4	5	8	7		

Jeu facile

5	3		7		
6			1	9	5
	9	8			6
8			6		3
4		8	3		1
7			2		6
	6			2	8
		4	1	9	5
		8		7	9

vous obtenez

9	6	3	1	7	4	2	5	8
1	7	8	3	2	5	6	4	9
2	5	4	6	8	9	7	3	1
8	2	1	4	3	7	5	9	6
4	9	6	8	5	2	3	1	7
7	3	5	9	6	1	8	2	4
5	8	9	7	1	3	4	6	2
3	1	7	2	4	6	9	8	5
6	4	2	5	9	8	1	7	3

Jeu difficile

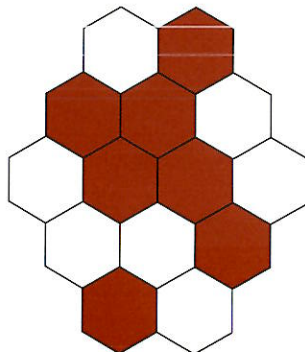
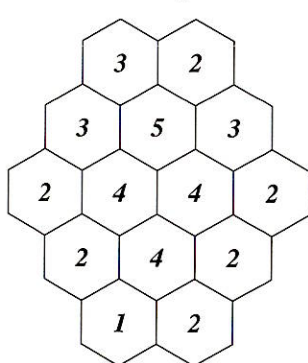
5		8	9	4	2			6
			1					
			5	7	3			
2	7		6		9			
3	4						1	5
			4		1		2	9
			3	9	4			
				8				
6			7	1	5	3		4

Ce nom bizarre provient du japonais « *Suji wa dokushin ni kagaru* » phrase qui peut à peu près se traduire par « les nombres doivent apparaître une seule fois ». Le jeu, parti de New York dans les années septante, est occupé à faire le tour du monde. (Pour plus d'informations, consulter le site [sudoku.sourceforge.net](http://sudoku.sourceforge.net)).

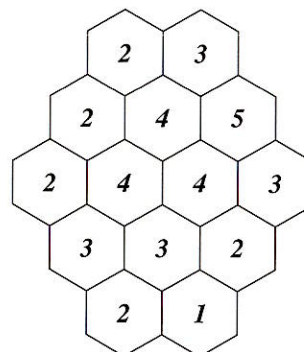
## 2. Les hexagones

Coloriez certains des hexagones des figures suivantes de façon que tout hexagone ait autant de voisins coloriés que le nombre indiqué en son centre. Attention : **tout hexagone est considéré comme voisin de lui-même !**

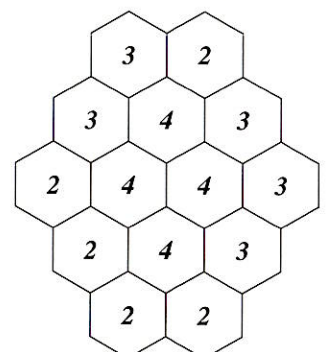
Exemple



Jeu facile

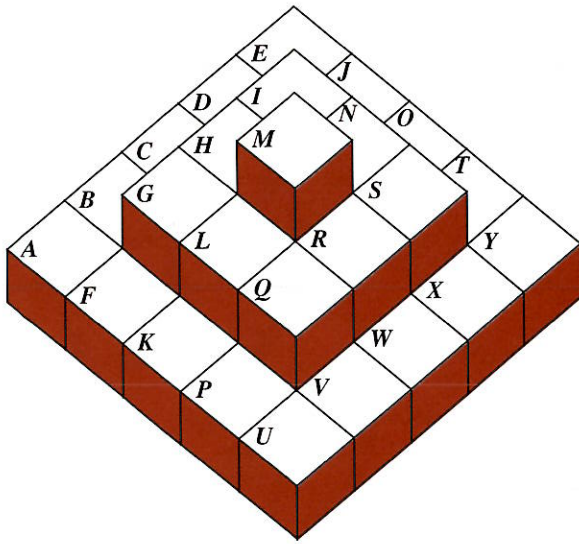


Jeu difficile





### 3. Pyramide magique<sup>(1)</sup>



Sur la pyramide, remplacer les lettres  $A, B, \dots, Y$  par les nombres de 8 à 32 (chacun étant utilisé exactement une fois) en respectant les consignes suivantes :

- En enfonçant d'un niveau le « cube  $M$  » on obtient un carré magique  $3 \times 3$  de somme 60
- et en enfonçant d'un niveau ce carré magique  $3 \times 3$  on obtient un carré magique  $5 \times 5$  de somme 100.
- Les nombres premiers se trouvent en  $F, G, J, Q, S, W$  et  $X$ .
- Les carrés parfaits sont en  $C, H$  et  $T$ .
- Les cubes parfaits sont en  $B$  et  $D$ .
- Les palindromes sont en  $J$  et  $N$ .
- Les puissances de 2 sont en  $B, H$  et  $V$ .
- Les multiples de 5 sont en  $A, M, P, T$  et  $Y$ .
- $W$  et  $X$  s'écrivent avec les mêmes chiffres permutés, de même que  $I$  et  $K$ .
- Les nombres impairs se trouvent en  $C, D, F, G, I, J, P, Q, S, T, W$  et  $X$ .
- Les multiples de 3 sont en  $A, C, D, I, K, L, P$  et  $R$ .

			63
			30
			192
180	18	112	

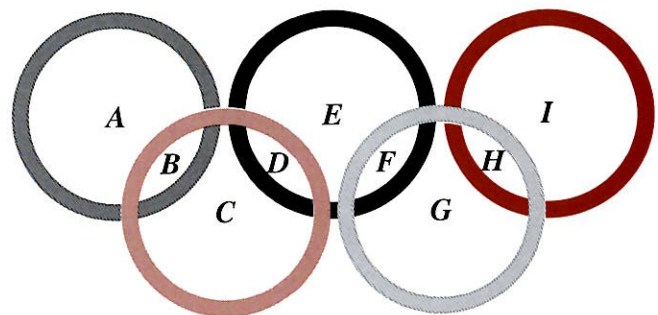
B. Le problème est analogue mais chacune des neuf cases contient un binôme  $ax + b$  (avec  $a$  et  $b$  entiers, éventuellement nuls).

			$4x^2 - 4x$
			$6x^3 + 3x^2$
			$4x^2 + 2x - 2$
$4x^3 - x$	$8x$	$3x^3 - 3x$	

### 4. Produits croisés

A. Les nombres de 1 à 9 occupent les neuf cases d'un carré  $3 \times 3$ . On donne les produits des trois nombres à droite de chaque ligne et en dessous de chaque colonne. Trouver le contenu des neuf cases.

### 5. Anneaux Olympiques<sup>(1)</sup>



Remplacer les lettres  $A, B, \dots, I$  par les nombres de 1 à 9 (chacun étant utilisé exactement une fois) de manière à ce que chaque anneau donne la même somme.



## 6. Anneaux olympiques et fantômes

En translatant, symétrisant, ... un fantôme, nous avons dessiné cinq « frises fantômes »

A.



B.



C.



D.



E.



Voici cinq « frises olympiques » dessinées de manières analogues aux précédentes :

1.



2.



3.



4.



5.



En tenant compte des processus de constructions, associer chaque « frise olympique » à une « frise fantôme ».

*Solutions des jeux page 24*

(<sup>1</sup>) Extrait de *Mathematical Pie*, N°165, 2005.



# Rallye problèmes

Nicole Miewis

Les vainqueurs ex-aequo du rallye problèmes de Math-Jeunes senior 2004-2005 sont Thomas Ra-delet (5<sup>e</sup>), Caroline Dubois (6<sup>e</sup>) et Lionel Di Pierdomenico (6<sup>e</sup>), tous trois de l'Athénée Vauban à Charleroi. **Toutes nos félicitations à ces amateurs de problèmes.**

Le rallye problèmes 2005-2006 comportera deux étapes publiées dans les numéros 111 et 112 de notre revue. À chaque étape, sept problèmes seront proposés à votre sagacité. Certains problèmes rapporteront plus de points que d'autres car ils demandent un raisonnement un peu plus long. Pour entrer en ligne de compte pour le classement final, nous vous demandons d'en résoudre au moins quatre à chaque étape. À vous de trouver une bonne stratégie de choix, en fonction du nombre de points rapportés par chaque problème, mais aussi de la confiance que vous avez en vos réponses. Quoi qu'il en soit, n'oubliez pas que *l'essentiel, c'est de participer.*

Rédigez les solutions des différents problèmes sur des feuilles séparées. La réponse finale ne suffit pas : expliquez et justifiez soigneusement vos solutions. Par contre vous pouvez nous envoyer des solutions partielles, par exemple si vous n'avez résolu qu'une partie du problème. Sur chaque feuille, indiquez vos nom, prénom, âge, adresse personnelle, classe et école.

Votre participation à la première étape doit être envoyée à N. MIÉWIS, avenue de Péville, 150, 4030 Grivegnée, munie de la mention « Concours Rallye Math-Jeunes Senior » pour le 15 décembre 2005. Les meilleures réponses seront publiées dans votre *Math-Jeunes*.

## 1. La chèvre (acte 1)

(5 points)

Un paysan prudent avait attaché sa chèvre par deux chaînes de cinq mètres à deux piquets plantés à cinq mètres l'un de l'autre. Quelle est la surface de la prairie que la malheureuse enchaînée pouvait brouter ?

## 2. La chèvre (acte 2)

(5 points)

Notre pauvre chèvre n'a pas fini d'en voir avec son brave (?) paysan de maître. Ce dernier possède également une prairie carrée dont la surface vaut 1 hectare. Il entend attacher sa chèvre à un piquet au milieu d'un côté de la prairie, par une corde d'une longueur telle que la chèvre pourra brouter exactement la moitié de la surface de la prairie. Quelle est (**à un mètre près**) la longueur de la corde ?

## 3. Le cauchemar

(5 points)

Un mordu de football, très affecté par la défaite de « son » équipe dormait d'un sommeil agité, et faisait un atroce cauchemar. Il se voyait une pauvre balle toute menue et l'adversaire un énorme ballon. Ce dernier le poursuivait dans toute sa chambre et il ne devait son salut qu'en se faisant tout petit, serré contre le mur et le parquet.

Au fait, si le ballon avait un diamètre de 30 cm, quel pouvait être, au maximum, le diamètre de la petite balle pour ne pas être écrasée sauvagement ?





#### 4. Le portefeuille

(5 points)

Le portefeuille du professeur a disparu et l'auteur du vol ne peut être que l'un des 5 élèves : Louise, Jean, Pierre, Albert ou Marie. Voici les réponses qu'ils firent au surveillant général chargé de l'enquête.

**Louise :** je n'ai pas pris le portefeuille et d'ailleurs je n'ai jamais rien volé de ma vie ; pour moi, c'est Albert le coupable.

**Jean :** mon père gagne assez d'argent et je ne manque de rien, je n'ai pas pris le portefeuille et puis interrogez Marie, elle sait qui est coupable.

**Pierre :** je n'ai pas le portefeuille ; c'est à l'école que j'ai connu Marie mais c'est Albert le coupable.

**Albert :** je ne suis pas coupable ; Louise ment quand elle dit que c'est moi ; la preuve c'est que c'est Marie qui a volé.

**Marie :** je n'ai pas pris le portefeuille ; Pierre qui me connaît depuis toujours, peut répondre de moi ; c'est Jean le coupable.

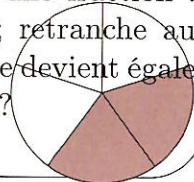
Plus tard, chacun(e) des cinq suspect(e)s reconnut que deux seulement des trois renseignements qu'il (elle) avait fournis étaient exacts. Qui est le coupable ?



#### 5. Histoire de fractions

(5 points)

Ajoute 1 aux deux termes d'une fraction : elle devient alors égale à  $\frac{222}{333}$  ; retranche au contraire 1 aux deux termes, elle devient égale à  $\frac{333}{666}$ . Quelle est cette fraction ?



#### 6. Les dominos

(5 points)

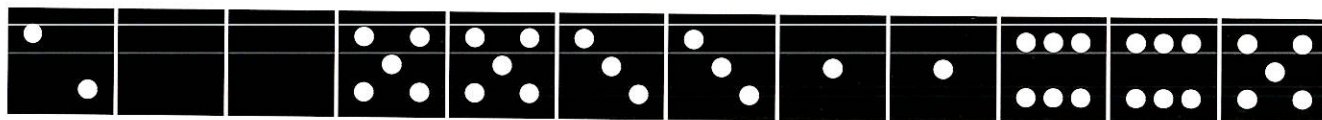
Un jeu de dominos est composé de 28 pièces. Sur chacune d'entre elles figure deux numéros choisis parmi les 7 premiers nombres naturels : de 0 à 6. Un même symbole peut apparaître deux fois. Si le choix des numéros se faisait parmi les entiers de 0 à 2005, de combien de pièces serait composé le jeu de domino ?

Un domino « standard » est un morceau de bois de peuplier (bois très léger de densité 0,35) de forme parallélépipédique dont les côtés sont de 8 mm, 24 mm et 48 mm. Ma camionnette possède un volume utile de  $2\text{ m} \times 2\text{ m } 50 \times 3\text{ m } 75$  et est limitée à une charge de 3,5 tonnes. Pourrais-je y transporter mon nouveau jeu ?

#### 7. 2005

(5 points)

Évaluer la somme  $S = [\sqrt{1}] + [\sqrt{2}] + \dots + [\sqrt{2005}]$  où  $[\sqrt{n}]$  désigne la partie entière de  $\sqrt{n}$ , c'est-à-dire le plus grand entier inférieur ou égal à  $\sqrt{n}$ .



Solutions des problèmes de la rubrique *Olympiades mathématiques*

C C E A D B C C A B

Éditeurs responsables et rédacteurs en chef :

– pour *Math-Jeunes* : G. NOËL, Rue de la Culée, 86, 6927 Restaigne

– pour *Math-Jeunes Junior* : A. PATERNOTTRE, Rue du Moulin 78, 7300 Boussu

© SBPMef Toute reproduction partielle ou totale est interdite sans autorisation.



# Math-Jeunes et Math-Jeunes Junior

Revues trimestrielles publiées par la  
Société Belge des Professeurs de Mathématique d'expression française

Secrétariat : M.-C. Carruana, SBPMef, rue de la Halle 15, 7000 Mons, ☎ 32-(0)65-373729,  
GSM : 0473-973808, e-mail : sbpm@sbpm.be, URL : <http://www.sbpn.be>.

## Math-Jeunes

**Comité de Rédaction :** C. Festraets, N. Lambelin, J. Mie-  
wis, N. Mewis, G. Noël, Y. Noël-Roch, F. Pourbaix, C. Ran-  
dour, P. Tilleuil, A. Tilleuil-Lepoutre, S. Trompler, N. Vande-  
nabeele, C. Villers.

## Math-Jeunes Junior

**Comité de Rédaction :** F. Drouin, C. Festraets, B. Hon-  
claire, G. Laloux, G. Noël, Y. Noël-Roch, A. Paternotte,  
F. Pourbaix, S. Trompler, N. Vandenabeele, C. Villers.  
**Mise en page et dactylographie :** M.-C. Carruana

## Tarifs

Trois numéros de chaque revue sont publiés annuellement. Pour connaître la liste des anciens numéros encore disponibles et leur prix selon votre localisation, adressez-vous au secrétariat.

Les abonnements destinés aux élèves de l'enseignement secondaire sont de préférence pris par l'intermédiaire d'un professeur. Effectuez vos paiements :

- ✉ pour la Belgique : par virement au Compte n°000-0728014-29 de S.B.P.M.e.f., rue de la Halle 15, 7000 Mons.
- ✉ pour la France : via l'APMEP (voir le Bulletin Vert).
- ✉ pour les autres pays : par virement international au Compte IBAN BE26 0000 7280 1429, BIC BPOTBEB1 de la SBP-Mef. Les personnes résidant à l'étranger qui ne peuvent faire un virement à ce compte sont invitées à envoyer un mandat postal international. Tout chèque bancaire non encaissable en Belgique devra être majoré de 12,50 € pour frais.

Abonnements groupés (5 exemplaires au moins)			
	Une des deux revues		Les deux revues
Belgique	4 €		8 €
Abonnements individuels			
	Une des deux revues		Les deux revues
Belgique	6 €		12 €
France (par APMEP)	8 €		16 €
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Europe	18 €	20 €	24 €
Autres pays	19 €	22 €	28 €

Non prior : ☐, Prior : ☐

Anciens numéros :

Avant 2004-2005 : 0,25 € par numéro + frais d'expédition

Année 2004-2005 : 0,50 € par numéro + frais d'expédition

Frais d'expédition : consulter le secrétariat

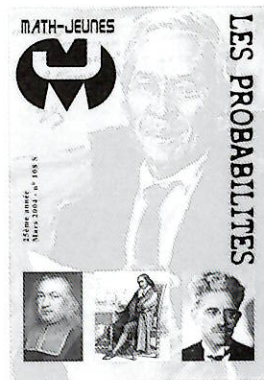
## Complétez votre collection de Math-Jeunes.



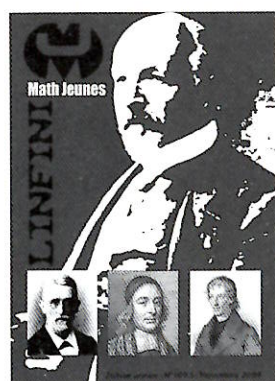
N°106  
L'algèbre



N°107  
Les courbes



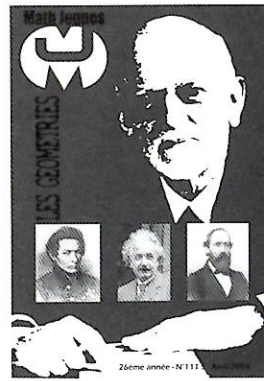
N°108  
Les probabilités



N°109  
L'infini



N°110  
Le codage



N°111  
Les géométries

Depuis le n°106, chaque numéro constitue une introduction à un thème particulier. Prix : 0,75 € pour les trois numéros 106 à 108 ; 1,5 € pour les trois numéros 109 à 111, frais d'expédition en sus.

Sont également toujours disponibles : les numéros 50 à 105 (sauf 61, 68 et 86). Les sommaires sont accessibles à l'adresse [www.sbpn.be/mj2.htm](http://www.sbpn.be/mj2.htm)



**Math-Jeunes**

Périodique trimestriel  
15, rue de la Halle – 7000 Mons  
Bureau de dépôt 7000 Mons 1

Responsable de l'édition: G. NOËL  
Rue de la Culée, 86 – 6927 Resteigne  
Bureau de dépôt: Mons 1

Autorisation de fermeture  
Sluitings toelating

7000 Mons 1  
5/156

Belgique - België  
P.P.  
7000 Mons 1  
5/124

Périodique - Peut être ouvert pour contrôle postal

Réservé à la poste

Inconnu

Refusé

Décédé

Adresse insuffisante

N'habite plus à l'adresse  
indiquée