collection textes de référence – Lycée (LEGT)
Document d'accompagnement des programmes

Mathématiques

Cycle terminal de la série ST2S

Outils pour la mise en œuvre des programmes 2007

Ministère de l'Éducation nationale
Direction générale de l'enseignement scolaire

Centre national de documentation pédagogique

Auteurs du document :

Gérard BONNEVAL	professeur de mathématiques, académie de Dijon
Alain MACÉ	IA-IPR de mathématiques, académie de Rouen
Magali ROUX	professeur de mathématiques, académie de Versailles
Magali BARALÉ	professeur de mathématiques, académie de Versailles
René MERCKOFFER	IA-IPR de mathématiques, académie de Versailles

Suivi éditorial : Corinne Paradas

Secrétariat d'édition : Nicolas Gouny

Maquette : A.M.G.

© CNDP, septembre 2007

Sommaire

Introduction

Ce document est destiné aux enseignants : il explicite et détaille les attendus du programme tout en proposant des pistes de réflexion pour permettre aux professeurs soit d'aborder certains points par des exemples variés, soit d'enrichir leurs connaissances personnelles.
La formation en mathématiques dans la série ST2S est conçue pour favoriser la poursuite d'études supérieures dans le domaine des sciences sanitaires et sociales ou à l'entrée dans la vie professionnelle et à la poursuite d'études au niveau bac + 2. La prise en compte de la diversité des parcours antérieurs des élèves est essentielle.
L'enseignement des mathématiques doit être relié à celui des autres disciplines afin, d'une part, de fournir les outils permettant de suivre avec profit les autres enseignements et d'autre part, de proposer des situations issues d'autres champs disciplinaires.
Le cadre et le vocabulaire théoriques doivent rester modestes, mais suffisamment efficaces pour répondre aux besoins mathématiques des autres disciplines et aussi pour jouer un rôle moteur lors des activités pluridisciplinaires.
Dans cette série, peut-être encore plus que dans d'autres, les élèves se sentent souvent en échec en mathématiques, notamment les élèves ayant suivi la voie professionnelle. Aussi, il est important de leur redonner confiance et de donner du sens aux mathématiques pratiquées.
Pour cela, on pourra par exemple travailler avec des exemples simples et notamment bâtis avec des données issues de domaines biologiques, socio-économiques, historiques et géographiques…
Les approches numériques, qui facilitent la compréhension des notions mathématiques, doivent tenir une large place. Les élèves doivent utiliser une calculatrice graphique ainsi que l'ordinateur.
Les activités graphiques doivent tenir une place importante. Elles développent qualité de soin et de précision, pertinence des interprétations. Elles mettent l'accent sur des travaux combinant un savoir-faire manuel, un appel à l'intuition et une réflexion théorique.
L'étude des statistiques et des probabilités doit constituer un moment important de la formation des élèves. En statistique, la lecture et la réalisation de tableaux et graphiques ont fait l'objet d'activités au collège et en seconde. Des situations adaptées doivent conduire les élèves à la pratique de la démarche statistique en tirant parti des possibilités offertes par les calculatrices et les outils informatiques. Les nouveaux résumés statistiques introduits doivent permettre aux élèves une approche plus approfondie de leur culture statistique et de ses applications.
En probabilités, le programme est une première initiation. L'objectif est d'entraîner les élèves à décrire des expériences aléatoires simples et à calculer des probabilités. Il s'agit d'éviter tout développement théorique.
On introduira la notion de probabilité, en s'appuyant sur la notion de fluctuation d'échantillonnage mise en évidence par la simulation abordée en classe de seconde. La notion de probabilité conditionnelle introduite en classe terminale prolonge naturellement celle de fréquence conditionnelle. Les notions de conditionnement et d'indépendance sont à même de faciliter l'analyse de résultats d'une enquête et favorisent l'esprit critique des élèves.

Les activités interdisciplinaires vont permettre aux élèves de connaître puis d'étudier des faits de société posant des questions sanitaires ou sociales en conjuguant des approches scientifiques, juridiques, socio-économiques, historiques, politiques, géographiques, culturelles… Elles ont pour objectifs de :
– porter un regard critique croisé sur des questions sanitaires ou sociales grâce à l'interdisciplinarité ;
– aborder les réponses dans leur diversité par, entre autres, une approche concrète du champ de la santé et du social ;

– enrichir les méthodes de travail de l'élève, développer la prise d'initiatives et la capacité à travailler en équipe (par le biais de travaux de groupe) ;
– développer les compétences nécessaires à la poursuite d'études et en particulier à l'accès aux concours sanitaires et sociaux.

Les thèmes de la classe de première et les études de la classe terminale relèvent du champ de la santé et du social et ne sont pas nécessairement une déclinaison d'un point du programme. Les activités menées doivent mettre en œuvre des acquis construits à partir des programmes de sciences et techniques sanitaires et sociales, de biologie et physiopathologie humaines, de sciences physiques et chimiques et de mathématiques.
Chaque année, trois semaines sont consacrées à ces activités. Elles sont menées sur les horaires de sciences et techniques sanitaires et sociales, de biologie et physiopathologies humaines, de sciences physiques et chimiques et de mathématiques, puisque ces quatre programmes ont intégré trois semaines réservées à ces activités.

Information chiffrée

Commentaires sur le programme

Information chiffrée et suites numériques

Dans tout le programme, mais peut-être plus encore dans cette partie, on pourra s'efforcer de développer un regard critique de l'élève sur les énoncés qui lui sont proposés. Cette attitude lui sera profitable, aussi bien en tant que citoyen, que dans le cadre de sa vie professionnelle. Par exemple, s'il se destine à une carrière d'infirmier, on peut lire dans l'article 29 du décret 93-221 du 16 février 1993, repris ensuite en 2005, relatif aux règles professionnelles des infirmiers ou infirmières : « L'infirmier ou l'infirmière applique et respecte la prescription médicale écrite, datée et signée par le médecin prescripteur, ainsi que les protocoles thérapeutiques et de soin d'urgence que celui-ci a déterminés. Il vérifie et respecte la date de péremption et le mode d'emploi des produits ou matériels qu'il utilise. Il doit demander au médecin prescripteur un complément d'information chaque fois qu'il le juge utile, notamment s'il estime être insuffisamment éclairé. L'infirmier ou l'infirmière communique au médecin prescripteur toute information en sa possession susceptible de concourir à l'établissement du diagnostic ou de permettre une meilleure adaptation du traitement en fonction de l'état de santé du patient et de son évolution. »
On pourra par exemple proposer des exercices avec des données surabondantes ou des données fausses (en précisant ce dernier point dans l'énoncé)…

Les élèves ont déjà vu dans les classes antérieures, aussi bien en mathématiques que dans d'autres disciplines, les pourcentages, mais cette notion est, encore au niveau du cycle terminal de cette série, source de nombreuses confusions.
Par ailleurs, pour des raisons de commodité de lecture sont regroupées dans cette partie plusieurs points ayant trait à l'utilisation des « pourcentages », mais traiter ensemble ces différentes notions risquerait d'entretenir cette confusion.
Aussi, on veillera à distinguer clairement les points suivants :
– « proportions », ou « fréquences », ou « rapport d'une partie au tout », ou encore « pourcentages instantanés » ;
– « taux d'évolution », ou « pourcentage d'évolution », ou « variation relative » ;
– « approximation linéaire dans le cas de faibles pourcentages » ou « approximation affine dans le cas de faibles pourcentages ».

N.B.
– Plusieurs expressions sont proposées afin que l'élève fasse le lien avec ce qu'il peut entendre dans d'autres disciplines.
– Un travail important pourra par ailleurs consister à traduire des phrases françaises (par exemple tirées de médias ou d'autres documents issus de domaines biologiques, médicaux ou paramédicaux…) en phrases mathématiques et réciproquement.

Le langage courant et les médias utilisent le mot « pourcentage » avec des sens divers, ce qui entraîne bien souvent confusions et incompréhensions chez les élèves comme chez beaucoup d'adultes. Il suffit pour s'en convaincre de comparer les phrases suivantes :
– « Le pourcentage de filles est 25 % dans le lycée. »
– « Le nombre de filles a augmenté de 25 %. »
– « Le pourcentage de filles a augmenté de 25 %. »
– « Le pourcentage de filles a augmenté de 25 points. »

En soi, le pourcentage n'est rien d'autre qu'une façon d'écrire les nombres décimaux $\frac{z}{100}$ s'écrit z %. On l'utilise couramment pour écrire les fréquences et les taux d'évolution. Mais ce sont les propriétés des fréquences et des taux d'évolution qui doivent être explicitées. Traiter ensemble ces deux notions très différentes, sous prétexte qu'elles s'écrivent toutes deux sous forme de pourcentage, conduit à des acrobaties de langage qui ne font qu'entretenir la confusion. C'est pourquoi le programme a nettement séparé les deux sous chapitres « Fréquences » et « Taux d'évolution ».

Pistes de réflexion pour le professeur

Fréquence (ou proportion)

La fréquence (ou proportion) de A dans E peut éventuellement se noter $f_E(A)$, ce qui préparera la notation des fréquences conditionnelles et des probabilités. Elle a des propriétés simples, notamment :
– Si p est la proportion de A dans E, et p' la proportion de B dans E, et si A et B sont disjoints, alors $p + p'$ est la proportion de $A \cup B$ dans E.
– Si p est la proportion de A dans E, et p' la proportion de E dans F, alors pp' est la proportion de A dans F.
On note que ces deux propriétés n'ont rien à voir avec les pourcentages : elles restent vraies si p et p' sont écrits sous forme décimale ou fractionnaire. Un usage tenace consiste pourtant à les appeler respectivement « addition de pourcentages » et « pourcentage de pourcentage ».

Tableau de contingence

Vocabulaire
Population : ensemble sur lequel porte l'étude.
Individu : un élément de cet ensemble.
Échantillon : une partie de la population
Recensement : enquête qui porte sur toute la population.
Sondage : enquête qui porte sur une partie de la population.
Partition : Soit E un ensemble non vide, une partition de E est un ensemble de parties de E qui sont non vides, disjointes deux à deux et dont la réunion est égale à E.
Caractère : critère qui permet de partager la population en réalisant une partition de l'ensemble.

Tableau d'effectifs pour deux caractères
On note r le nombre de lignes (nombre de modalités du caractère A) et s le nombre de colonnes (nombre de modalités du caractère B).

B / A	b_1	…	b_j	…	b_s	Marge de A
a_1						$n_{1\bullet}$
…						
a_i			n_{ij}			$n_{i\bullet}$
…						
a_r						
Marge de B	$n_{\bullet 1}$		$n_{\bullet j}$		$n_{\bullet s}$	n

nij : i pour le numéro de la ligne, puis j pour celui de la colonne.
n_{ij} est le nombre d'individus qui possèdent la modalité a_i du caractère A et la modalité b_j du caractère B. Les effectifs lus à l'intérieur du tableau de contingence sont appelés les effectifs partiels.
$n_{i\bullet}$ est le nombre d'individus qui possèdent la modalité a_i du caractère A.

$n_{\bullet j}$ est le nombre d'individus qui possèdent la modalité b_j du caractère B.
Les effectifs lus dans les deux marges du tableau sont appelés effectifs marginaux.

Tableau des fréquences pour deux caractères

– Fréquences partielles : les fréquences partielles sont obtenues en faisant le rapport des effectifs partiels à l'effectif total. La proportion d'individus possédant à la fois la modalité a_i du caractère A et la modalité b_j du caractère B est notée f_{ij}, on a donc : $f_{ij} = \frac{n_{ij}}{n}$.

– Fréquences marginales : les fréquences marginales sont obtenues en faisant le rapport des effectifs marginaux à l'effectif total. La proportion d'individus possédant la modalité a_i du caractère A est notée : $f_i = \frac{n_i}{n}$.

– Fréquences conditionnelles : les fréquences conditionnelles sont obtenues en faisant le rapport de l'effectif partiel à l'effectif marginal. La proportion d'individus possédant la modalité a_i du caractère A sous condition qu'ils possèdent la modalité b_j du caractère B est notée $f_{i/j}$ et on a : $f_{i/j} = \frac{n_i}{n_{\bullet j}}$.

Tableau des fréquences partielles

B / A	b_1	…	b_j	…	b_s	Marge de A
a_1						$f_{1\bullet}$
…						
a_i			f_{ij}			$f_{i\bullet}$
…						
a_r						
Marge de B	$f_{\bullet 1}$		$f_{\bullet j}$		$f_{\bullet s}$	1

Ces notations ne sont pas à introduire dans un cours. Cependant, il est possible de demander un calcul de fréquence partielle (ou de pourcentage) ou de fréquence conditionnelle (ou de pourcentage) à partir d'une phrase clairement exprimée.

Vocabulaire

Tri à plat : on entend par cette expression un tri de la population selon un seul caractère.
Tri croisé : on trie la population suivant deux caractères.
Tableau de contingence : tableau des effectifs de la population triée selon deux caractères.

Exemple et contre-exemple

Cas 1 : tableau de contingence

Le tableau suivant donne les immatriculations, une certaine année, des véhicules d'une région.

B / A	Département 1	Département 2	Département 3	Département 4
Voitures neuves	9 724	15 939	8 703	10 859
Voitures d'occasion	28 460	43 794	29 290	30 320
Cars et autobus neufs	23	23	32	17
Cars et autobus d'occasion	26	43	51	43
Camionnettes, camions neufs	1 824	2 712	1 805	8 326
Camionnettes, camions d'occasion	5 107	8 499	5 004	5 239

Voici une liste non exhaustive de questions que l'on peut poser à partir de ce tableau :
– Que signifie le nombre 29 290 ?
– Dans le département 1, quel est le pourcentage des immatriculations de camionnettes et camions neufs par rapport au parc des véhicules de ce département 1 ?
– Dans quel département, le pourcentage d'immatriculations de camionnettes ou camions neufs est-il le plus faible ?
– Donner la répartition en pourcentages des immatriculations de véhicules dans le département 4.

Cas 2 : tableau qui n'est pas de contingence
L'individu considéré dans le tableau suivant est la commune.

B / A	Département 1	Département 2	Département 3	Département 4
École de musique	15	57	53	55
Musée	14	39	18	20
Bibliothèque	73	111	88	101

Les différentes modalités de la variable A (équipement culturel) ne permettent pas de faire une partition des communes (une commune peut à la fois avoir une école de musique, sa bibliothèque ainsi que son musée), même si les modalités de la variable B permettent une partition. Ce tableau n'est pas un tableau de contingence.

Taux d'évolution (ou variation relative)

L'expression variation relative qui désigne $\frac{y_2 - y_1}{y_1}$ vise à la distinguer de la variation absolue $y_2 - y_1$. Dans cette définition y_1 et y_2 sont deux nombres réels strictement positifs. En pratique, ce sont deux valeurs successives d'une même grandeur, mesurées avec la même unité. Par suite, la variation relative est sans unité, contrairement à la variation absolue. Une variation relative peut s'exprimer sous forme décimale, de fraction ou de pourcentage. Mais par convention, une variation exprimée en pourcentage est toujours une variation relative.
Le taux d'évolution a une propriété essentielle, qui découle immédiatement de la définition : dire que t est le taux d'évolution entre y_1 et y_2 équivaut à dire que $y_2 = y_1(1 + t)$. Autrement dit que $\frac{y_2}{y_1} = 1 + t$.

N.B. – Dans le programme de ST2S, comme par exemple dans celui de 1re L Math-Info, le nombre 1 + t est appelé « coefficient multiplicatif ». Dans celui de STG, il est appelé « coefficient multiplicateur » ou « multiplicateur ». Ces légères différences sont sans importance et on pourra utiliser n'importe laquelle de ces expressions.
Par ailleurs, comme il est écrit dans la brochure *Dé-chiffrer par les maths*, il est important de noter que le taux est un nombre. On a donc diverses écritures du même nombre : $t = \frac{5}{100} = 0{,}05$ qui se notera aussi 5 %. Dans ce cas, « 82 % de 1 200 » se traduit directement par 0,82 × 1 200. Le coefficient multiplicatif, dans le cas d'un pourcentage d'évolution, sera donc $(1 + t)$ et non pas, comme dans certains manuels, $1 + \frac{t}{100}$.

Approximation linéaire dans le cas de faibles pourcentages

Cette partie est à séparer nettement des parties précédentes. C'est une notion délicate, car après avoir montré précédemment que deux augmentations successives de 2 % ne correspondaient pas à une augmentation globale de 4 % (et donc casser des idées ancrées fortement dans l'esprit des élèves et de nombreux adultes !), cette partie du programme revient sur les approximations couramment utilisées.

– L'idéal serait bien évidemment d'étudier ces approximations dans le cadre de l'utilisation du nombre dérivé.
– Ou bien de se contenter d'une interprétation graphique en faisant tracer les courbes d'équation $y = (1 + t)^2$ et $y = 1 + 2t$ au voisinage de 0.
– Dans le cas contraire, on pourrait se contenter de travailler sur quelques exemples et montrer que les termes négligés dans $(1 + t)^n$ (avec $n = 2$ ou 3) sont peu importants pour des valeurs de t proche de 0.
– Quelle que soit la démarche choisie, on pourra utiliser avec profit un tableur afin de faire des comparaisons des résultats obtenus en faisant varier le taux et le nombre d'évolutions successives. On pourra lire avec profit la brochure *Dé-chiffrer par les maths*, pages 141 à 143.

Tableur

La difficulté dans l'utilisation d'un logiciel (tableur, logiciel de géométrie dynamique…) est le rôle laissé à l'élève :
– soit il est très guidé et son rôle se limite alors à recopier plus ou moins bien des suites d'instructions dont il ne comprend pas toujours le sens ;
– soit une plus grande latitude lui est laissée et il risque de perdre une grande partie de son temps à essayer de régler des problèmes « techniques » en se demandant sur quelle touche, ou combinaison de touches, il doit appuyer pour obtenir le résultat souhaité.
Aussi, pour une séance donnée, une démarche envisageable peut être la réalisation de deux activités successives et semblables, où pour la première activité, l'élève serait très guidé, puis pour la deuxième, il n'aurait que peu d'indications et devrait donc reconstruire un cheminement semblable à celui déjà effectué.
L'importance de la première étape diminuerait au fur et à mesure de l'avancée de l'année scolaire.
Par ailleurs, il semble important qu'à l'issue de chaque séance, une trace écrite soit demandée.
Pour ne pas multiplier les chapitres, ce paragraphe a été rattaché à l'information chiffrée. Mais son libellé indique bien que l'utilisation du tableur concerne aussi bien les suites que les statistiques et les fonctions.
L'édition d'une formule, la copie d'une formule, l'utilisation d'un adressage absolu ou relatif, la dénomination de cellules, la réalisation d'un graphique sont les compétences de base à développer ou à consolider dans le cadre d'activités de résolution de problèmes.
Les « règles et contraintes assez simples » évoquées dans la colonne des capacités attendues sont celles qui permettent :
– de définir les fonctions de référence ;
– de calculer les termes de suites arithmétiques ou géométriques (formules d'itération, formules explicites) ;
– d'étudier des séries statistiques (calculs de fréquences, de moyenne, de médiane, de déciles, de quartiles, d'écart type, de maximum, de minimum, etc.) ;
– d'effectuer des simulations (tirage aléatoire, partie entière, condition du type « si… alors » à utiliser mais non exigible en évaluation, dénombrement conditionnel, etc.).
On veillera à ce que les séances de travail sur tableur donnent lieu à une production écrite, faute de quoi l'expérience montre que les méthodes ne sont pas mémorisées.
Il importe également qu'elles donnent lieu à des questions ciblées lors des évaluations. Il s'agit de repérer certains concepts, notions et outils mathématiques mis en œuvre lors de l'utilisation d'un tableur (notamment les notions de variable, de fonction, de moyenne pondérée).
À partir d'exemples (budgets d'association, feuilles de remboursement de la sécurité sociale, feuilles de facturation, etc.), on s'attachera à comprendre comment se font les modifications de toutes les cellules de la feuille de calcul lorsqu'on change une donnée, une pondération ou une règle de calcul.
On ne parlera pas des tableaux théoriques ou dits de proportionnalité ; les commentaires sur les pourcentages des lignes (resp. des colonnes) se feront simplement à

partir des distributions de fréquences associées aux marges horizontales (resp. verticales). On pourra prendre comme exemple de référence l'étude de résultats d'élection (classification selon les régions ou les classes d'âge des votes à une élection où plusieurs candidats sont en présence).

Interprétation des proportions : attention aux erreurs possibles

Dans de nombreuses études on veut comparer deux ou plusieurs sous populations d'une population donnée au vu de la distribution dans chacune des sous populations d'une variable qualitative prenant un nombre fini de valeurs possibles. On calcule alors des proportions que l'on compare et on interprète les différences constatées. Mais cette interprétation comporte des pièges, on va en explorer certains d'entre eux à l'aide de quelques exemples.

L'effet Simpson ou le coup de la variable cachée

Le statisticien américain Simpson recherche si les universités ne pratiquent pas, dans leur sélection des candidats, une discrimination dont les femmes seraient les victimes. En effet, pour l'université X, on a les résultats suivants :

Université X	Hommes	Femmes	Total
Admis	534	113	647
Refusés	664	336	1 000
Total	1 198	449	1 647

Pour l'université X : parmi les hommes, 44,6 % d'admis, parmi les femmes, 25,2 % d'admis.
Il semble bien qu'une discrimination existe dont les femmes seraient les victimes.
On continue l'étude. L'université est composée de deux départements A et B ; que se passe-t-il dans chacun de ces départements ?
Pour le département A, on a les résultats suivants :

Département A	Hommes	Femmes	Total
Admis	512	89	601
Refusés	313	19	332
Total	825	108	933

Pour le département A : parmi les hommes, 62,1 % d'admis, parmi les femmes, 82,4 % d'admis.
Il est clair que le département A ne discrimine pas les femmes dans son recrutement.
Pour le département B, on trouve :

Département B	Hommes	Femmes	Total
Admis	22	24	46
Refusés	351	317	668
Total	373	341	714

Pour le département B : parmi les hommes, 5,9 % d'admis, parmi les femmes, 7,0 % d'admis.
Dans le département B, comme dans le département A, les procédures d'admission semblent plus favorables aux femmes qu'aux hommes.

Si chaque composante de l'université fait réussir davantage les femmes que les hommes parmi ses candidats ou candidates, pourquoi au niveau global les femmes semblent-elles désavantagées ? On observe que la variable « département » n'est pas prise en compte dans le premier tableau alors que les deux départements sont très différents.

B est beaucoup plus sélectif que A mais il attire davantage les femmes que A. Cherchons, selon le sexe, la répartition des candidatures.
Parmi les hommes : 31 % sont candidats en B.
Parmi les femmes : 76 % sont candidates en B.
Cela suffit à expliquer le paradoxe, mais il fallait penser à faire intervenir la variable « département ».
On trouve un paradoxe analogue quand on compare les réussites scolaires des enfants de nationalité française et ceux de nationalité étrangère. À catégorie sociale identique, les Français semblent mieux réussir, aussi, en terme de remédiation, on pense, pour les enfants de nationalité étrangère, à un renforcement de l'enseignement du français.
Mais si on compare les réussites, non seulement à catégorie sociale identique, mais aussi à même taille de fratrie, on observe que les enfants de nationalité étrangère réussissent mieux que les Français. La remédiation à faire se trouverait plutôt du côté des études dirigées que des cours de français, il est plus difficile de faire ses devoirs et d'apprendre ses leçons quand les frères et sœurs sont plus nombreux. Pour plus de détails consulter [1], page suivante.

Quand le poids des sous populations change

Il est une statistique que des ministres ont publiée et que les médias ont diffusée et commentée en commettant une grave erreur d'interprétation. Il s'agit de montrer que la sélection sociale par l'école s'est aggravée depuis 1945. Pour cela on s'intéresse aux deux populations de reçus dans les grandes écoles : Polytechnique (X), École normale supérieure (ENS), École nationale d'administration (ENA), Hautes études commerciales (HEC), en 1950 et en 1993, la variable étudiée est leur origine sociale. Pour les deux années la population concernée représente 0,12 % d'une classe d'âge. Parmi les reçus on distingue ceux originaires du milieu populaire et ceux originaires du milieu intellectuel défini comme ceux dont le père a un niveau d'études supérieur ou égal au baccalauréat. Le tableau suivant donne, en pourcentage, leur part respective pour les années de référence.

Reçus aux concours	En 1950	En 1993
Milieu populaire	25 %	9 %
Milieu intellectuel	60 %	80 %
Autres	15 %	11 %

Du point de vue quantitatif le verdict est sans appel, il y a moins de jeunes issus du milieu populaire dans les futures élites en 1993 qu'en 1950. Mais est-ce de la faute de l'école ? Ceux qui à la vue du tableau précédent, en déduisent un phénomène de recul de la démocratisation raisonnent comme si la composition sociale de la société française n'avait pas évolué entre les deux années, or il n'en n'est rien. Le tableau ci-dessous résume cette évolution.

Population française	En 1950	En 1993
Milieu populaire	80 %	60 %
Milieu intellectuel	5 %	20 %
Autres	15 %	20 %

On note que l'évolution est considérable.
Si on choisit au hasard un jeune dans l'une et l'autre des sous populations étudiées, comment a évolué sa probabilité de faire partie des reçus ? Le calcul est simple, on obtient les résultats suivants :

Probabilité d'être reçu × 10^4	En 1950	En 1993
Milieu populaire	3,75	1,8
Milieu intellectuel	144	48

La part plus importante prise par le milieu intellectuel au détriment du milieu populaire dans la société globale fait qu'il est plus difficile pour les jeunes des deux milieux

de faire partie des reçus. En gros deux fois plus difficile pour les jeunes du milieu populaire mais trois fois plus difficile pour ceux du milieu intellectuel. Contrairement aux apparences du premier tableau, l'école n'a pas rendu la tâche plus facile aux héritiers de la culture ; pire leur situation s'est davantage dégradée que celle des jeunes issus du milieu populaire. Dans ces conditions peut-on encore affirmer que l'école a aggravé la sélection sociale entre les deux dates ? Mais le raisonnement précédent est-il accessible au journaliste moyen ?
On peut rendre le phénomène que nous avons analysé plus intuitif en imaginant des situations extrêmes. Supposons qu'à une époque antérieure il existe une catégorie sociale, représentant une proportion infime de la population mais dont les enfants ont une probabilité très grande de rentrer dans l'une des très grandes écoles. Si à une date ultérieure cette catégorie croît alors cette probabilité va s'effondrer, sinon elle aurait davantage de reçus que le nombre total de reçus, ce qui est absurde.

Pour affiner la réflexion il faudrait modéliser autrement les inégalités sociales. Il existe une continuité quant à la proximité des cultures vécues dans les familles et celle de l'école, le regroupement de la population en trois catégories supposées homogènes n'en rend pas compte. Il est bien connu que la réussite scolaire des enfants d'ouvriers qualifiés est nettement supérieure à celle des enfants d'ouvriers dits spécialisés, pour ne prendre qu'un seul exemple.

Références

[1] P. Dutarte et J.-L.Piednoir, *Enseigner la statistique au lycée : des enjeux aux méthodes,* commission inter-IREM lycées technologiques.
On pourra également consulter :
[2] Les publications de l'association Pénombre.
[3] Le site Statistix : www.statistix.fr

Mathématiques et usage de l'informatique

Commentaires sur le programme

Dans le programme il est dit : « Il convient qu'en ce domaine les professeurs déterminent en chaque circonstance la stratégie d'utilisation la mieux adaptée afin de mettre l'outil informatique au service des apprentissages. »
Comme il est hors de question de faire un « cours » de tableur ou de grapheur, l'introduction à doses homéopathiques des différentes fonctions de celui-ci, au gré des besoins, semble s'imposer.

Pistes de réflexion pour le professeur

Thèmes possibles	Tableur
Tableau à double entrée	Introduction feuille de calculs : **cellule, ligne, colonne, format de cellule**
La notion de variable (mode de génération de suites numériques ou introduction des fonctions numériques)	**Formule** écrite dans une cellule et opérations de base, **référence absolue, référence relative, poignée de recopie**
Représentations graphiques (fonctions, statistiques)	Assistant graphique (types de graphiques)
Les fonctions mathématiques du programme	Arrondi, somme, racine carrée
Les fonctions statistiques du programme	Moyenne, écart type, aléa, NB.SI…

Apprendre à transcrire des données dans un tableur, puis en tirer des informations
– Apports nutritionnels pour les adolescents (annexe 1).
– Remboursement de prestations de la Sécurité sociale (annexe 2).
– Résultats d'analyses médicales (glycémie et cholestérol) (annexe 3).

Annexe 1 – Les apports nutritionnels du repas de midi conseillés pour les adolescents

Si une proportion entre les glucides, les protides et les lipides doit être respectée, il convient de bien répartir les différentes catégories d'aliments. Ainsi il faut prévoir :
– quatre fois plus de glucides « lents » que de glucides « rapides » ;
– autant de protides d'animaux que de protides végétaux ;
– deux fois plus de graisses végétales saturées que de graisses animales insaturées.

	Apports nutritionnels	Moyenne
Énergie	3 600 kJ	
Protéines	25,5 à 27 g	26,25
Lipides	30 à 31,5 g	30,75
Glucides	115,5 à 117 g	116,25

En utilisant les valeurs moyennes dans le cas d'intervalle remplir le tableau suivant :

Protides	animales	
	végétales	
Glucide	lents	
	rapides	
Graisses	végétales saturées	
	animales insaturées	

Annexe 2 – Remboursement de prestations de la Sécurité sociale

	A	B	C	D	E	F	G	H
1					Sécurité sociale		Complémentaire	
2	Soins du …	Concernant	Montant des soins	Base de remboursement	%	Montant	%	Montant
3	01/03/07		25	20	70			
4	02/03/07		22,96	22,96	65			
5								
6								
7								
8	Viré le	Montant						
9	01/05/07							

En supposant que l'assurance complémentaire permette un remboursement complet de la base de remboursement, compléter cette feuille de relevé de prestations par deux méthodes différentes.

En remplissant dans cet ordre les cellules F3, G3, H3, puis les colonnes F, G, H, et enfin la cellule B9.

Annexe 3 – Résultats d'analyses médicales

	A	B	C	D	E	F
1						
2	Monsieur	Dupont				
3						
4			Valeurs souhaitables			Commentaires
5			min	max		
6	Glycémie en g/L	1,01	0,7	1,1		OK
7	Acide urique en mg/L	45	35	70		OK
8	Triglycérides en g/L	0,39		1,5		OK
9	Cholestérol total en g/L	2,31		2,2		Attention
10	Cholestérol HDL en g/L	0,55		0,45		Attention
11	Rapport Total/HDL	4,2		5		OK
12	Cholestérol LDL en g/L	1,68		1,6		Attention

a) Remplir la cellule D10 en tenant compte de la cellule A2, sachant que les valeurs souhaitables ont un maximum de 0,55 pour une femme et 0,45 pour un homme.
b) Remplir la cellule D11.
c) Établir une feuille automatisée qui dans la colonne F répondra :
– « OK » si les résultats sont dans la fourchette attendue ;
– « Attention » sinon.

Évolution du nombre de logements construits à Toulouse entre 1988 et 1992

La séquence qui suit, élaborée à partir d'un document de l'académie de Toulouse, fait appel à un nombre faible de données, pour lesquelles l'usage d'un tableur n'est peut-être pas indispensable, les calculs pouvant être menés à l'aide d'une calculatrice. C'est le principe de l'activité qui est intéressant, car il met en jeu de nombreux questionnements sur les pourcentages. Il peut être reproduit sur des données plus nombreuses issues de l'économie ou d'autres domaines (données qu'il serait alors difficile de reproduire intégralement sur un document papier).

On donne le nombre de logements individuels construits à Toulouse entre 1988 et 1992 (source : document mairie de Toulouse). L'objectif est d'étudier l'évolution (en nombre et en pourcentage) de ces constructions.

	A	B	C	D	E	F
1			*Construction de logements*			
2						
3	*Années*	**1988**	**1989**	**1990**	**1991**	**1992**
4	***Individuels purs***	151	160	133	129	131
5	**Evolution en nombre**	-----------------				
6	**Evolution en pourcentage**	-----------------				
7						
8	***Individuels groupés***	210	124	186	174	62
9	**Evolution en nombre**	-----------------				
10	**Evolution en pourcentage**	-----------------				
11						
12	***Ensemble***					
13	**Evolution en nombre**	-----------------				
14	**Evolution en pourcentage**	-----------------				

N.B. – Les logements individuels purs désignent les maisons isolées et les logements individuels groupés les maisons dans des lotissements.

a) Calculer, à l'aide du tableur, le nombre total de logements individuels construits chaque année.

Quelle est la formule à saisir en B12 ?

Comment faire effectuer les autres calculs par le tableur ?

b) Que penser de la phrase : « Le nombre de logements individuels à Toulouse a diminué de 1989 à 1990 » ?

c) Calculer à la main :

– l'évolution en nombre du nombre de logements individuels purs construits entre 1988 et 1989 ;

– l'évolution en pourcentage du nombre de logements individuels purs construits entre 1988 et 1989 ;

– l'évolution en nombre du nombre de logements individuels groupés construits entre 1988 et 1989 ;

– l'évolution en pourcentage du nombre de logements individuels groupés construits entre 1988 et 1989 ;

– l'évolution en nombre du nombre total de logements individuels construits entre 1988 et 1989 ;

– l'évolution en pourcentage du nombre total de logements individuels construits entre 1988 et 1989. Quelle remarque pouvez-vous faire ?

d) À l'aide du tableur, faire calculer les contenus de toutes les cellules C5 à F6, C9 à F10, C13 à F14 :

– Quelle formule saisir en C5 ? En C6 ?

– Comment calculer les contenus des cellules D5 à F6 ?

– Comment calculer les contenus des cellules C9 à F10 ?

– Comment calculer les contenus des cellules C13 à F14 ?

e) On note $f(n)$ le nombre total de logements individuels construits l'année n.

– À l'aide du tableur, construire les points A_n de coordonnées $(n, f(n))$.

– Calculer à la main le coefficient directeur de la droite $(A_{1988}\, A_{1989})$.

– À quel nombre déjà calculé correspond-il ?
– De manière plus générale, quelle est l'expression du coefficient directeur de la droite (A_n A_{n+1}) ?
Quelle est la signification de ce nombre dans le tableau ?

Information chiffrée et solveur

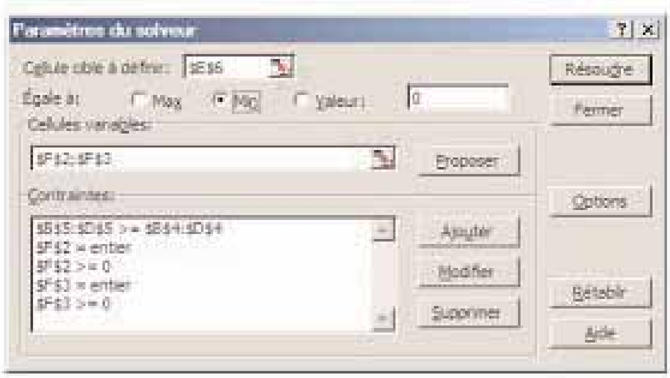
	A	B	C	D	E	F
1		Heures de fictions françaises	Heures de fictions étrangères	Heures de reportages	Coût unitaire	Quantité optimale
2	x	90	120	40	20000	13
3	y	90	60	100	24000	7
4	Nombre d'heures à diffuser	1800	1800	1200		
5	Nombre d'heures achetées	1800	1980	1220		
6				Coût minimal	428000	

=B2*F2+B3*F3

=C2*F2+C3*F3

=D2*F2+D3*F3

Les quantités optimales (cellules F2 et F3) sont obtenues à l'aide du SOLVEUR programmé de la manière suivante :



Fonctions numériques

Commentaires sur le programme

Le programme recommande de poursuivre le travail sur la notion de fonction mis en place en seconde, en s'appuyant sur la pratique de problèmes, à partir de situations issues des sciences biologiques et physiques et de la vie économique et sociale. Cela suggère de travailler en liaison avec les collègues enseignant ces disciplines, et de choisir aussi souvent que possible t et non x comme variable en référence au temps. D'autres noms de la variable pourront être présentés dans des situations en lien avec les sciences physiques et chimiques, biologie et physiopathologie humaine, sciences et techniques de la santé et du social. L'exploitation systématique des interprétations graphiques des notions et des résultats étudiés se retrouve au cœur des travaux des élèves, en particulier dans le cas où les méthodes algébriques connues des élèves sont inopérantes.

Sans faire un chapitre de révisions, il s'agit de consolider les acquis de la classe de seconde, en particulier pour tout ce qui concerne les représentations graphiques de fonctions et leurs interprétations en termes de résolutions graphiques ou de sens de variation.

Les différentes notions pouvant être revues au détour d'exemples tirés des programmes de biologie ou de sciences physiques sont les suivantes :

– image et antécédent(s) d'un nombre ;
– sens de variation sur un intervalle ;
– résolution graphique d'équations du type $f(t) = k$ ou $f(t) = g(t)$.

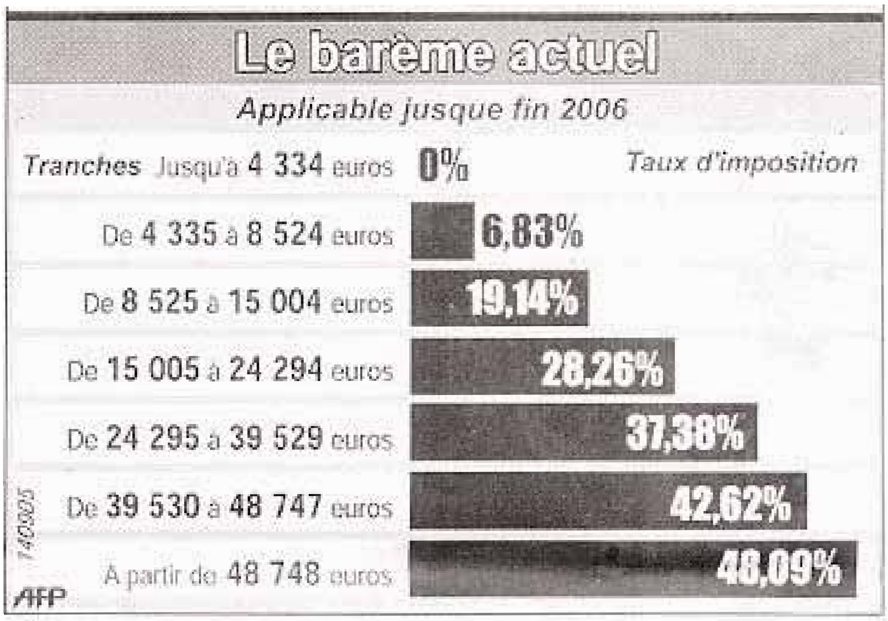
En classe de première

L'objectif essentiel est de permettre une bonne vision de l'apport des fonctions et de leurs représentations graphiques, ce qui, en général, est loin d'être acquis en fin de seconde ou de BEP. Les contenus des programmes de ces deux niveaux sont très proches sur la notion de fonction. Il faudra cependant tenir compte de l'absence dans les programmes de BEP Carrières sanitaires et sociales de la notion de fonction sinus et de fonction cosinus. Il sera intéressant de réactiver les connaissances acquises au collège sur le sinus et le cosinus d'un angle aigu dans un triangle rectangle à l'occasion d'exercices : leur utilisation interviendra en particulier en optique (loi de Descartes : $n_1 \sin i_1 = n_2 \sin i_2$ et travail d'une force : $F \times l \times \cos a$ dans le programme de première des sciences physiques et chimiques) mais non évaluable en mathématiques.

La notion capitale de sens de variation est souvent mal comprise. Des phrases simples peuvent permettre d'assimiler la définition : « Une fonction croissante conserve le sens des inégalités ; une fonction décroissante renverse le sens des inégalités. » L'interprétation graphique est indispensable, mais n'explique pas pour quelle raison cette notion est importante. Comme en seconde, il faut multiplier les exemples où une fonction relie deux grandeurs variant soit dans le même sens, soit en sens opposés :

– L'impôt est une fonction *continue* et croissante du revenu ainsi que le revenu net disponible.

Impôt sur le revenu : la nouvelle donne



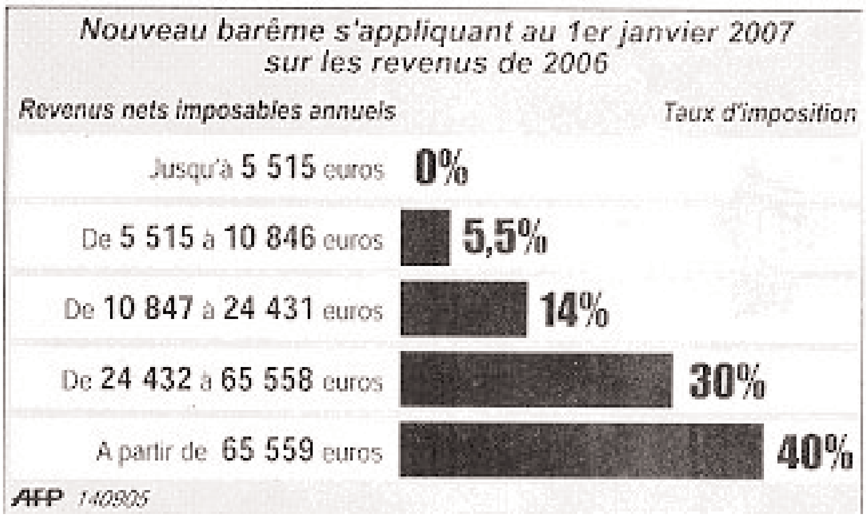
Le barème actuel

Applicable jusque fin 2006

Tranches	Taux d'imposition
Jusqu'à 4 334 euros	0%
De 4 335 à 8 524 euros	6,83%
De 8 525 à 15 004 euros	19,14%
De 15 005 à 24 294 euros	28,26%
De 24 295 à 39 529 euros	37,38%
De 39 530 à 48 747 euros	42,62%
A partir de 48 748 euros	48,09%

AFP 140905

Le ministre des Finances Thierry Breton a dévoilé hier les détails d'une vaste réforme fiscale qui vise à réduire les taux et le nombre de tranches de l'impôt sur le revenu.



Nouveau barème s'appliquant au 1er janvier 2007 sur les revenus de 2006

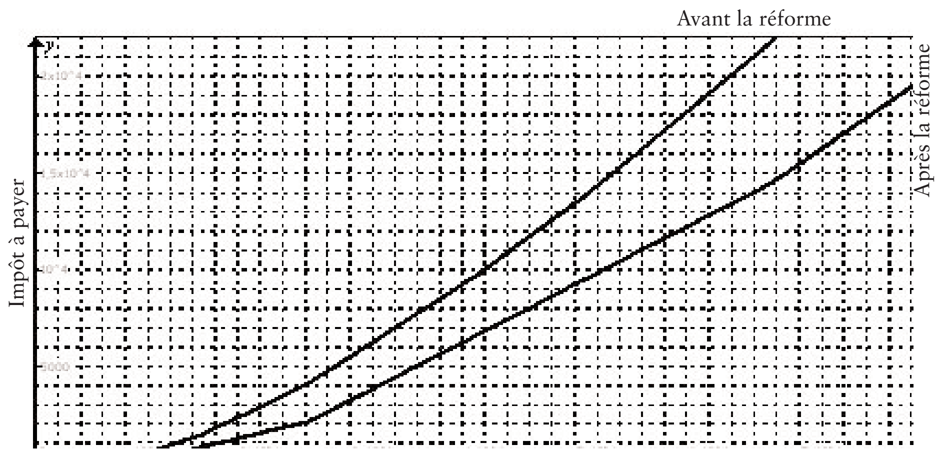
Revenus nets imposables annuels	Taux d'imposition
Jusqu'à 5 515 euros	0%
De 5 515 à 10 846 euros	5,5%
De 10 847 à 24 431 euros	14%
De 24 432 à 65 558 euros	30%
A partir de 65 559 euros	40%

AFP 140905

On se propose de représenter les variations du montant y en fonction du revenu net imposable annuel x.

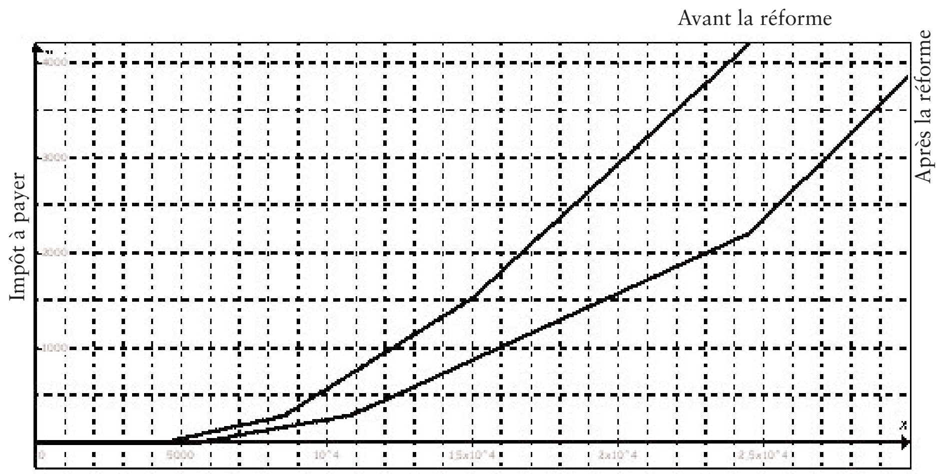
Droite	Coefficient directeur	Passant par le point	Et par le point	Équation
D_1	0,0683	(4 335 ; 0)	(8 525 ; 286)	$y = 0{,}0683x - 296$
D_2	0,1914	(8 525 ; 286)	(15 005 ; 1 526)	$y = 0{,}1914x - 1\,345$
D_3	0,2826	(15 005 ; 1 526)	(24 295 ; 4 151)	$y = 0{,}2826x - 2\,714$
D_4	0,3738	(24 295 ; 4 151)	(39 530 ; 9846)	$y = 0{,}3738x - 4\,930$
D_5	0,4262	(39 530 ; 9846)	(48 748 ; 13 775)	$y = 0{,}4262x - 7\,001$
D_6	0,4809	(48 748 ; 13 775)		$y = 0{,}4809x - 9\,668$
D_7		(5 515 ; 0)	(10 847 ; 293)	$y = 0{,}055x - 303$
D_8		(10 847 ; 293)	(24 432 ; 2 195)	$y = 0{,}14x - 1\,225$
D_9		(24 432 ; 2195)	(65 559 ; 14 533)	$y = 0{,}30x - 5\,134$
D_{10}	0,40	(65 559 ; 14 533)		$y = 0{,}40x - 11\,690$

Lorsqu'on représente ces fonctions sur des intervalles bien définis, on obtient la représentation graphique suivante :



Revenu net imposable

Et effectuant un zoom sur la figure précédente, l'on obtient :



Revenu net imposable

Ceci représente une fonction affine par intervalles. Changer de tranche, c'est changer de fonction affine. Le taux d'imposition y est plus élevé. Est-ce pour autant que le changement entraîne un saut significatif dans l'impôt ?

Un rapide calcul pour un revenu en 2005 de 15 004 € et un revenu de 15 005 € (changement de tranche) donne, pour le premier cas, un impôt de 1 526,77 € et, dans le second cas, un impôt de soit moins de 1 526,41 € soit une baisse de 0,36 € pour une augmentation de revenu de 1 €. Il n'y a donc pas de saut, la fonction I est continue. C'est l'objectif de la constante soustractive ajoutée dans chaque fonction affine. L'impression subjective de discontinuité provient du fait que le taux change brusquement : la pente de la droite change brusquement créant un point anguleux. On dit que la fonction n'est pas dérivable. Une deuxième idée couramment répandue est que le saut de tranche peut entraîner une perte de revenu réel. Le revenu disponible, qui correspond au revenu brut duquel on soustrait les impôts, est représenté par une fonction continue sur R. C'est de nouveau une fonction affine par intervalles. Les pentes de tous les segments de droites ou demi-droites sont positives et la fonction $R_1 = R - I \times R$ est continue donc elle est aussi une fonction croissante : plus on gagne d'argent, plus il en reste après impôts.

Autres exemples possibles :
– La taille d'un enfant est fonction croissante de son âge.
– La concentration d'un produit après son injection dans le sang est fonction décroissante du temps.
– Le temps mis pour parcourir une distance donnée est une fonction décroissante de la vitesse.

Fonctions de référence

Les fonctions affines ainsi que les fonctions carré et inverse ont été vues en seconde et en BEP. On leur adjoint en première les fonctions cube et racine carrée. Elles constituent, avec les fonctions logarithme décimal et exponentielle de base *a* qui seront abordées en classe terminale, les éléments qui permettront de fabriquer les fonctions usuelles.
Le souci permanent chez l'enseignant d'un changement de cadre entre les registres graphiques et algébriques doit permettre de donner du sens aux situations proposées.
L'utilisation des calculatrices graphiques, des tableurs et des logiciels de géométrie dynamique, tant en démonstration dans la classe avec un vidéo-projecteur, ou en utilisation individuelle par les élèves, a profondément modifié la problématique de l'étude des fonctions. La courbe se retrouve ainsi à constituer fréquemment la première étape de cette étude. Cela suppose pour l'élève des compétences spécifiques :
– sur calculatrice ou sur un logiciel adapté à l'ordinateur, savoir « entrer » une fonction, la tabuler, choisir une fenêtre adaptée pour la représenter, savoir obtenir la pente (ou coefficient directeur) en un point donné de cette représentation ;
– sur tableur, savoir « entrer » une formule, la recopier, construire un graphique.
Ces compétences n'ont rien d'immédiat et demandent par conséquent un apprentissage trop souvent négligé. Cela représente aussi pour l'enseignant des difficultés concrètes :
– pour la calculatrice, le problème de la diversité des modèles ;
– pour le tableur, l'accès régulier en effectif adapté à une salle d'informatique fiable n'est pas toujours possible.
Une deuxième étape consiste à interpréter les informations de la calculatrice ou du tableur : images, antécédents, signe, sens de variations, *extrema*, positions relatives de deux courbes. Ces notions ont été introduites en seconde et en BEP mais sont loin d'être maîtrisées en première. Passer du cadre graphique au cadre numérique et inversement (comparer une table de valeurs et une courbe, des coefficients directeurs de tangentes et les nombres dérivés, placer des points, lire des images et des antécédents, lire un sens de variation, repérer un *extremum*…) permet la maturation de ces notions.
Il importe aussi de leur donner du sens en les ancrant dans des situations concrètes : un élève de première ST2S doit pouvoir expliquer pour quelles raisons la quantité d'un médicament dans le sang après une injection est une fonction du temps nécessairement décroissante ; l'utilité d'obtenir ou de comparer les vitesses d'absorption d'un produit par l'organisme ; de comparer les fonctions donnant la vitesse d'augmentation d'une population de bactéries en milieu renouvelé ou non renouvelé…

Une fois ces compétences acquises, une troisième étape consiste pour l'élève à comprendre que ce qu'il voit à l'écran permet des conjectures, qui seront autant que possible démontrées par exemple *via* un changement de cadre comme le calcul algébrique (dans des cas où les calculs n'exigent pas trop de technicité).

Nombre dérivé et tangente

On se contente en première ST2S d'introduire le nombre dérivé, dont la compréhension demande une longue maturation. La fonction dérivée et l'observation du signe du nombre dérivé seront traitées en classe terminale. Le passage du nombre $f'(a)$ à la fonction f' présente en effet des obstacles, notamment l'absence de tracé de courbe pour cette nouvelle fonction. Néanmoins, le terrain sera préparé par les formules concernant le nombre dérivé pour les fonctions de référence. Le travail se fera donc essentiellement par des approches graphiques amenant à la lecture de coefficients directeurs de tangentes à une courbe, ou à la construction de la tangente en un point à une courbe d'équation $y = f(t)$ (voir fiches 52 et 48 du site Euler).

– Le nombre dérivé en 1 de la fonction carré est 2, ce qui fournit une formule d'approximation locale $(1 + t)^2 \approx 1 + 2t$ pour t proche de 0. Cette formule a une application aux petits taux d'évolution : augmenter deux fois de suite de 1 % revient sensiblement à augmenter de 2 %.

– Le nombre dérivé en 1 de la fonction inverse est – 1, ce qui permet d'écrire $\frac{1}{1+t} \approx 1 - t$ pour t proche de 0 : on peut en déduire par exemple que le taux réciproque de 1 % est sensiblement – 1 %.

– Le nombre dérivé en 1 de la fonction racine carrée est $\frac{1}{2}$, ce qui permet d'écrire $\sqrt{1+t} \approx 1 + \frac{t}{2}$ pour t proche de 0. On peut en déduire par exemple que si après deux variations relatives identiques le taux global est 4 %, c'est que chaque variation relative était sensiblement 2 %.

Sur chaque exemple numérique, il est essentiel de comparer la valeur exacte et des valeurs approchées de différents niveaux de précision. Dans le premier cas, on peut même le faire de façon littérale en développant $(1 + t)^2$.

Certains logiciels comme *Geogebra* permettent d'introduire de façon visuelle le nombre dérivé en « zoomant » autour d'un point ou de faire déplacer une tangente le long d'une courbe en affichant son coefficient directeur (voir fiche 48 du site Euler).
Les approches faisant intervenir la notion de vitesse (comme pour des concentrations d'un produit dans le sang permettant d'obtenir la vitesse d'absorption par l'organisme) sont aussi à utiliser. La recherche hors contexte d'une équation d'une tangente n'est pas un objectif du programme.
Un exemple de vitesse instantanée, donc de nombre dérivé, est donné couramment par l'indicateur de vitesse d'une automobile :

En ce qui concerne les fonctions de référence carré, inverse, cube, racine carrée, on admet les formules $2a, -\frac{1}{a^2}$, $3a^2$ et $\frac{1}{2\sqrt{a}}$ donnant le nombre dérivé en a (voir fiche 47 du site Euler). L'utilisation d'un tableur peut permettre de les tester pour quelques valeurs de a ; la construction et l'observation de la droite passant par le point A $(a\,;f(a))$ et ayant ce coefficient directeur permet de renforcer le sens de la notion de tangente en A à la courbe.

En classe terminale : fonction dérivée

Le nombre dérivé a été défini en classe de première et les formules qui permettent de le calculer ont été admises pour les fonctions de référence. En classe terminale, la notion de fonction dérivée permet de réinvestir ce travail, mais il est important de prendre en compte que cette notion n'est pas un des objectifs principaux de la formation. Il s'agit d'avoir à sa disposition un outil supplémentaire pour l'étude des variations de certains phénomènes, mais l'enseignant doit se limiter à des objectifs extrêmement modestes. En dehors des cas de fonctions dérivées d'une somme de fonctions de référence ou du produit par un réel, la fonction dérivée sera fournie.

Pistes de réflexion pour le professeur

Liens mathématiques / biologie

Exemple 1 : indice de masse corporelle

Avant l'entrée des enfants à l'école primaire, les médecins et infirmières du ministère de l'Éducation nationale réalisent un bilan de santé et mesurent la taille (en mètres) et la masse (en kilogramme) de chaque enfant. Ces deux paramètres permettent d'obtenir l'indice de masse corporelle (IMC) indicateur d'une éventuelle surcharge pondérale.

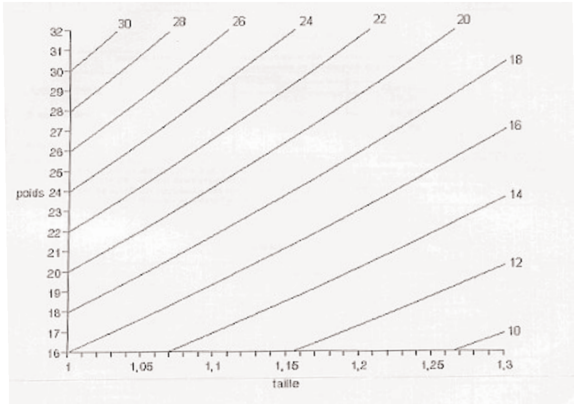
L'indice de masse corporelle (IMC) se calcule à l'aide de la relation suivante : IMC = $\frac{P}{T^2}$

où P désigne le poids de l'enfant en kilogrammes et T sa taille en mètres.
Sur un échantillon de 15 garçons de six ans, on a relevé le poids et la taille de façon à calculer l'IMC de ces enfants.

Poids (P)	15	18	19	20	22	23	23	
Taille (T)	1,05	1,05	1,09	1,10	1,02	1,17	1,15	
IMC	13,61	16,33		16,53	21,15	16,80	17,39	
Poids (P)	23	24	24	25	25	26	27	27
Taille (T)	1,16	1,20	1,20	1,25	1,30	1,20	1,24	1,30
IMC	17,09	16,67	16,67	16,00	14,79	18,06	17,56	15,98

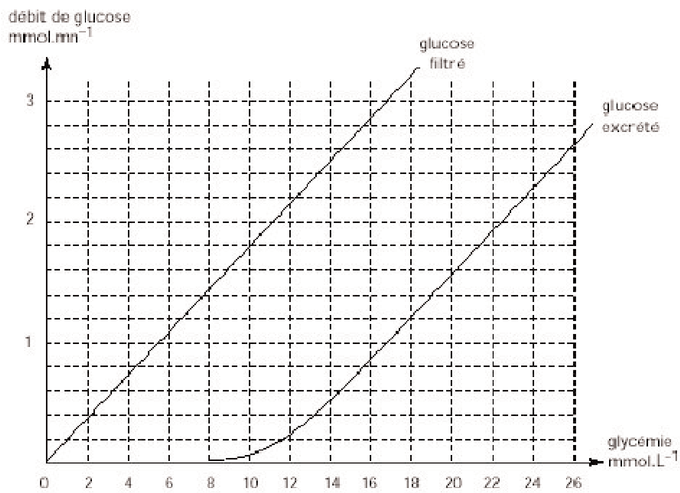
Calculer l'IMC d'un enfant de 19 kilogrammes mesurant 1,09 mètre.
Le graphique ci-après représente différentes courbes de niveaux de la surface S donnant l'IMC en fonction du poids (entre 12 et 32 kilogrammes) et de la taille (entre 1 mètre et 1,3 mètre) de l'enfant.
– Donner une valeur approchée de l'IMC d'un enfant pesant 19 kilogrammes et mesurant 1,09 mètre.
– Donner une valeur approchée du poids d'un enfant dont l'IMC vaut 20 et mesurant 1,22 mètre.
– Dans quel intervalle se situe le poids d'un enfant mesurant 1,25 mètre dont l'IMC est compris entre 12 et 16 ?

Courbes de niveaux de IMC = 10 à IMC = 30 de la surface S



Exemple 2 : glycémie

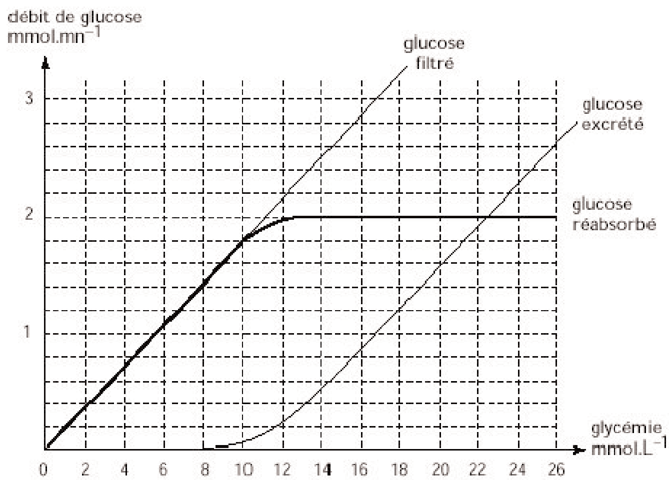
Le document suivant représente le débit de glucose filtré et le débit de glucose excrété dans l'urine en fonction de la concentration en glucose plasmatique.



a) Trouver la valeur de la glycémie à partir de laquelle il y a présence de glucose dans les urines.

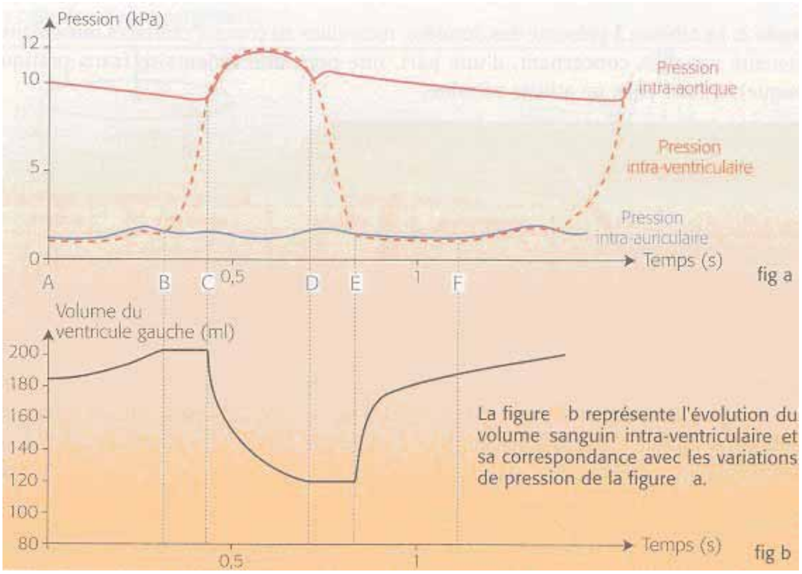
Le glucose commence à être excrété lorsque la glycémie est supérieure à 9 mmol.L^{-1} comme le montre le document. Il y aura donc présence de glucose dans les urines quand la glycémie dépassera cette valeur.

b) Tracer sur le document, la courbe représentant le débit de glucose réabsorbé en fonction de la glycémie.



c) Définir et trouver à partir de la courbe la capacité maximale de réabsorption rénale du glucose.
Le glucose est entièrement réabsorbé lorsque la glycémie est inférieure à 9 mmol.L^{-1}. Au delà, le glucose commence à être excrété dans l'urine, les capacités de réabsorption rénale étant dépassées : c'est la capacité maximale de réabsorption du glucose. La courbe du glucose réabsorbé montre que la capacité maximale de réabsorption est égale à 2 mmol.min^{-1}.
d) Préciser le mécanisme selon lequel se fait le transport trans-membranaire du glucose.
Le glucose est réabsorbé par un mécanisme de transport actif, c'est-à-dire consommateur d'énergie, dans le tube contourné proximal. Il s'agit d'un cotransport avec les ions sodium.

Exemple 3 : variation de la pression artérielle en fonction du temps

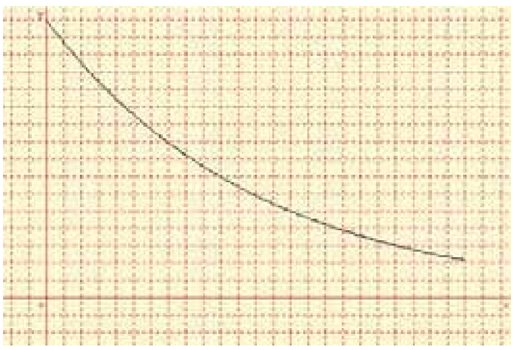


Le schéma de la figure b engendre le tableau de variation ci-dessous.

Temps (s)	*A* *0*	*B* *0,3*	*C* *0,4*	*D* *0,7*	*E* *0,8*	*F* *1,1*
Volume du ventricule gauche (mL)	185	200	200	120	120	200

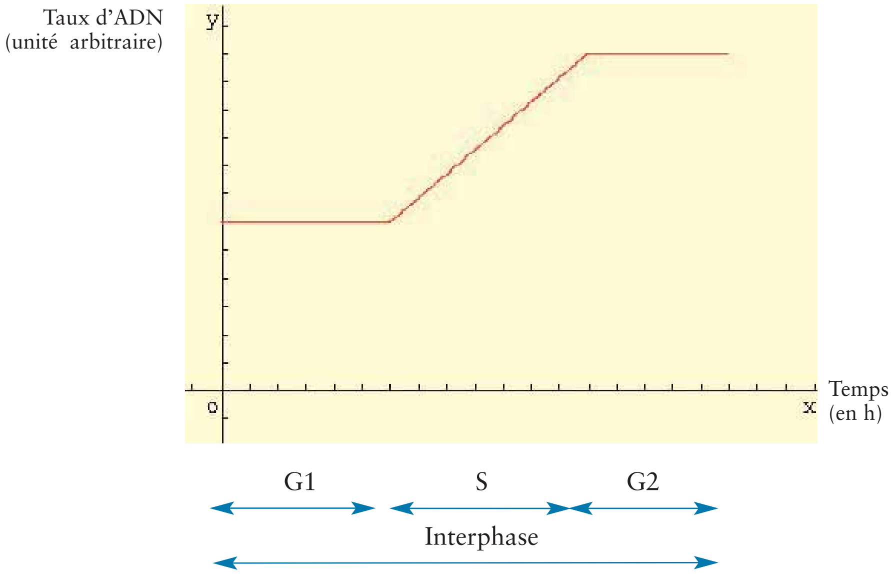
Exemple 4 : quantité de médicament présente dans le sang après injection

On injecte à un malade une dose de 2 centimètres cubes d'un certain médicament M. La quantité de médicament présente dans le sang du malade pendant les 24 heures suivant l'injection est donnée par la courbe ci-dessous :



On peut s'aider des données ci-dessus afin de légender le graphique puis :
– parler des variations de cette quantité de médicament ;
– déterminer graphiquement le temps écoulé après l'injection pour que la quantité de médicament présente dans le sang soit la moitié de la dose injectée (qui était de 2 cm^3) ;
– donner une approximation, à l'unité près (ce qui permet de revoir la notion de valeur approchée) de la quantité restant dans le sang au bout de 24 h par rapport à la dose injectée.

Exemple 5 : variation du taux d'ADN au cours du cycle cellulaire



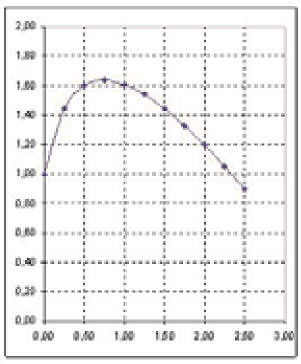
L'interphase se divise en trois phases de durées inégales : deux phases G1 et G2 séparées par une phase de synthèse d'ADN (phase S). On peut décrire en termes mathématiques ces différentes phases :
– taux d'ADN constant ou croissant ;
– taux à différentes heures ;
– durée de chacune des phases ;
– plage horaire ou durée de la synthèse d'ADN (période pendant laquelle la cellule a doublé son taux d'ADN.
On peut aussi sur cet exemple chercher à déterminer l'expression de la fonction sur l'intervalle [6 ; 12], puisqu'il s'agit, sur cet intervalle, d'une fonction affine.

Exemple 6 : taux de glycémie et nombre dérivé

On étudie la glycémie d'une personne observée après ingestion de sirop de glucose.
Des analyses effectuées tous les quarts d'heure et pendant 2 heures et demie ont donné les résultats suivants :

Temps en heures	0	0,25	0,5	0,75	1	1,25	1,5	1,75	2	2,25	2,5
Glycémie en $g.L^{-1}$	1	1,44	1,60	1,64	1,61	1,54	1,45	1,33	1,20	1,05	0,90

On note t le temps exprimé en heures et $f(t)$ la glycémie à l'instant t, exprimée en $g.L^{-1}$.
À partir de ce tableau de valeurs, on a représenté la fonction $t \mapsto f(t)$.
Déterminer graphiquement le sens de variation de la fonction f.
On constate que la glycémie augmente jusqu'à atteindre son maximum à l'instant 0,75, puis elle diminue.
On se propose d'étudier la vitesse de régulation du glucose.



– Vitesse moyenne : les données numériques permettent de calculer la vitesse moyenne entre les instants $t - 0{,}25$ et $t + 0{,}25$, pour tout réel t appartenant à $\{0{,}25 ; 0{,}5 ; \ldots ; 2{,}25\}$.

Cette vitesse moyenne est égale à $\dfrac{f(t+0{,}25)-f(t-0{,}25)}{(t+0{,}25)-(t-0{,}25)} = \dfrac{f(t+0{,}25)-f(t-0{,}25)}{0{,}5}$.

On note $V_1(t)$ cette vitesse. On obtient les résultats suivants :

t	0,25	0,5	0,75	1	1,25	1,5	1,75	2	2,25
$V_1(t)$	1,2	0,39	0,02	– 0,19	– 0,331	– 0,42	– 0,5	– 0,56	– 0,6

Pourquoi certaines vitesses sont-elles négatives ?
Interprétation graphique : coefficients directeurs des sécantes.
Il semble que la vitesse de la réaction soit toujours décroissante.

– On veut préciser la vitesse de la réaction aux alentours du premier quart d'heure, c'est-à-dire lorsque $t = 0{,}25$.
a) On dispose d'une nouvelle série de mesures qui ont été réalisées à intervalles de temps fixes.

Temps en heures	0,15	0,25	0,35
Glycémie en g.L^{-1}	1,32	1,44	1,5255

Calculer la vitesse moyenne $V_2\,(0{,}25)$ entre les instants 0,15 et 0,35.
Interpréter graphiquement.

b) On dispose de deux autres séries de mesures.

Temps en heures	0,2	0,25	0,3
Glycémie en g.L^{-1}	1,3878	1,44	1,4885

Temps en heures	0,24	0,25	0,26
Glycémie en g.L^{-1}	1,4329	1,44	1,4529

Calculer les vitesses moyennes $V_3(0{,}25)$ entre les instants 0,2 et 0,3 et $V_4(0{,}25)$ entre les instants 0,24 et 0,26.
Observation des résultats :

h	0,25	0,1	0,05	0,01
Vitesse moyenne entre $0{,}25 - h$ et $0{,}25 + h$	1,1972	1,02733	1,0067	1,00027

En donnant à h des valeurs de plus en plus petites, on définit des vitesses moyennes sur des intervalles d'amplitude de plus en plus petite, centrés en 0,25.

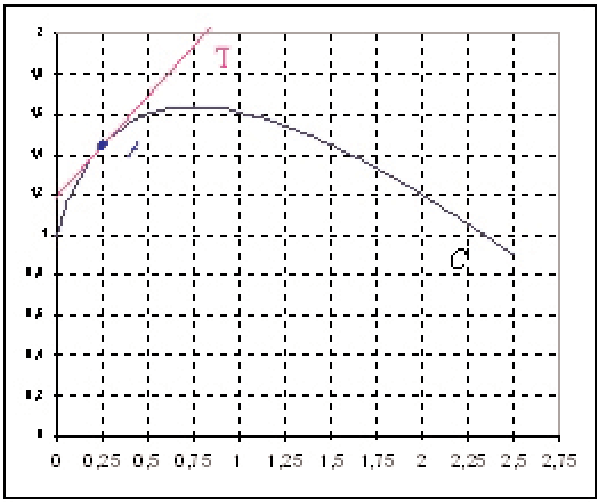
h	0,0050			Vitesse moyenne
0,2450	1,4381	0,2450	1,4381	1,00006667
0,2500	1,4431	0,2550	1,4481	
0,2550	1,4481			

h	0,0010			
0,2490	1,4421	0,2490	1,4421	1,00000267
0,2500	1,4431	0,2510	1,4441	
0,2510	1,4441			

h	0,0005			
0,2495 0,2500	1,4426 1,4431	0,2495 0,2505	1,4426 1,4437	1,00000069
<0,2505	1,4437			

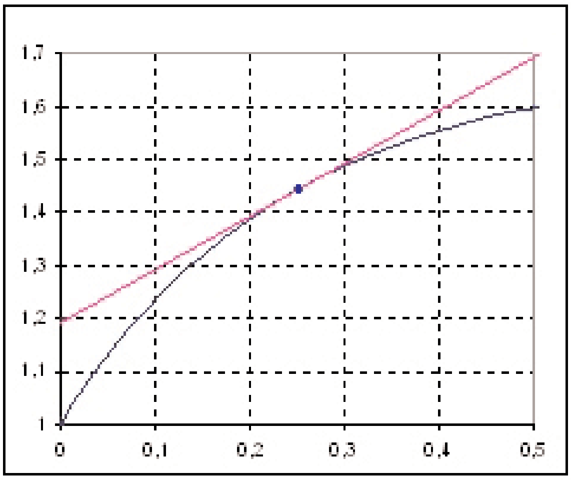
Ces vitesses moyennes varient de moins en moins et semblent se stabiliser autour de 1. On dira que la vitesse de la réaction à l'instant 0,25, c'est-à-dire après un quart d'heure, est égale à 1.

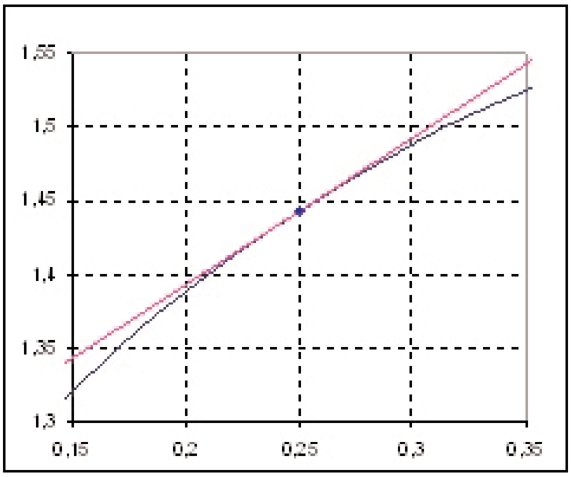
– Interprétations graphiques : dans chaque cas, la vitesse moyenne entre les instants $0{,}25 - h$ et $0{,}25 + h$ est le coefficient directeur de la droite passant par les points de coordonnées $(0{,}25 - h\,;\, f(0{,}25 - h))$ et $(0{,}25 + h\,;\, f(0{,}25 + h))$.
On a tracé la courbe C représentative de la fonction f et la droite T passant par le point A d'abscisse 0,25 de cette courbe de coefficient directeur égal à 1. Que constate-t-on ?



La droite T est appelée la *tangente* en A à la courbe.
Le coefficient directeur de cette droite est la vitesse instantanée de la réaction de régulation de la glycémie.

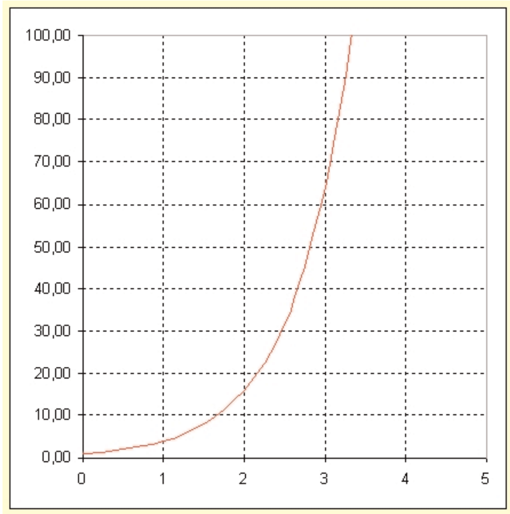
– Nombre dérivé : des agrandissements permettent de préciser la position de la courbe par rapport à la tangente.



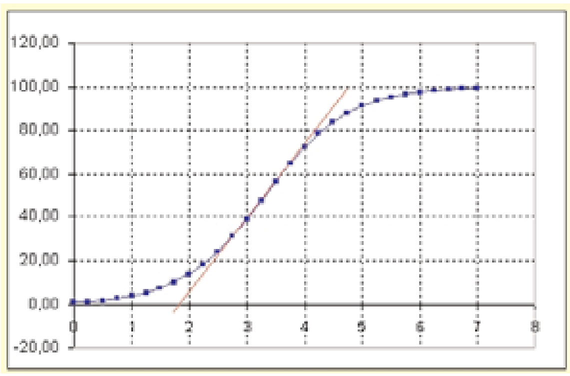


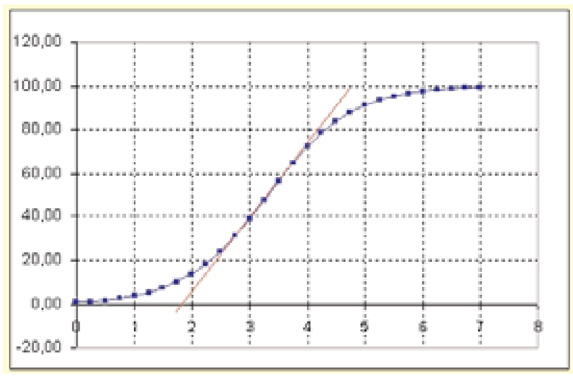
Exemple 7 : étude d'une population de bactéries

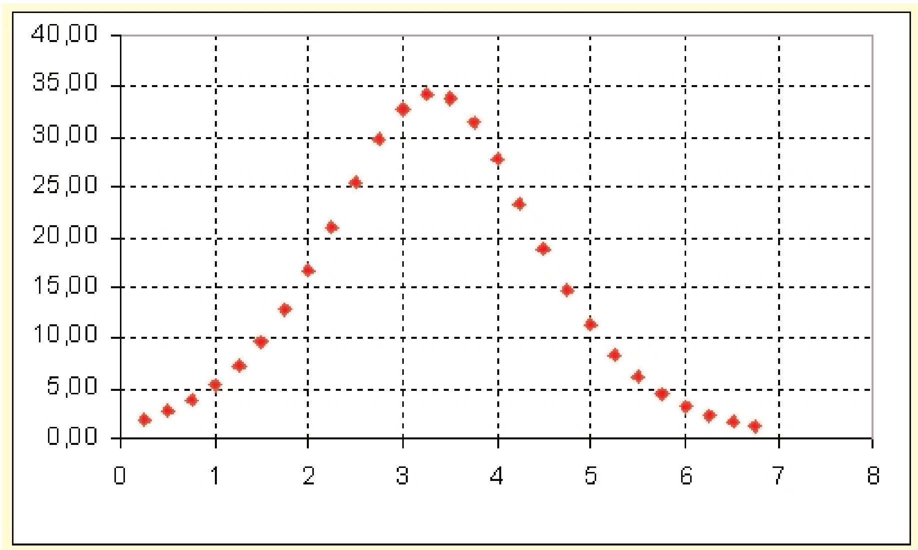
a) En milieu renouvelé : la vitesse augmente indéfiniment. La courbe est toujours au-dessus de sa tangente.



b) En milieu non renouvelé : la vitesse passe par un maximum.
Il semble que cette vitesse maximale soit atteinte pour t = 3,5 avec une valeur égale à 34,25.
On peut tracer la droite passant par le point d'abscisse 3,5, de coefficient directeur égal à 34,25 et observer que cette droite traverse la courbe.



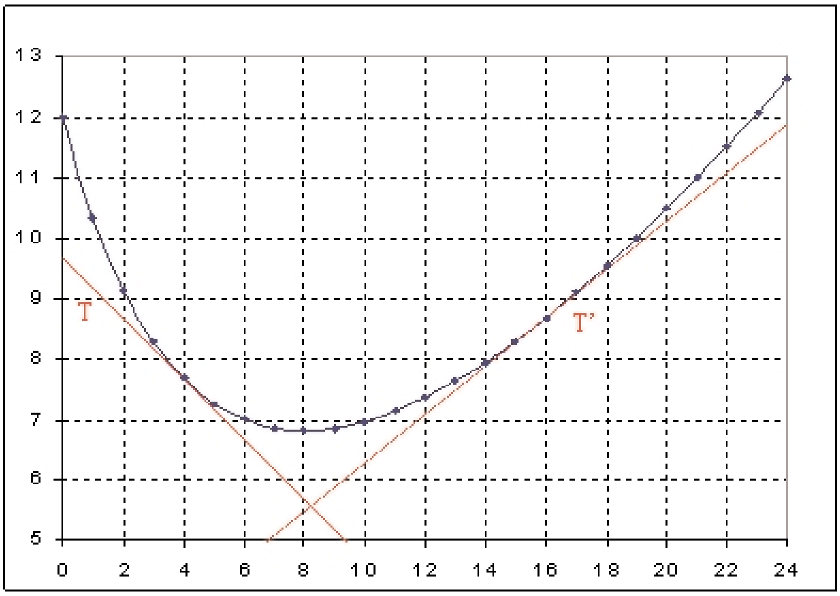




Exemple 8 : étude du taux d'anticorps chez un nourrisson de moins de deux ans

Des analyses ont permis d'établir la courbe suivante, représentant le taux d'anticorps dans le sang d'un nourrisson, de sa naissance jusqu'à l'âge de deux ans.

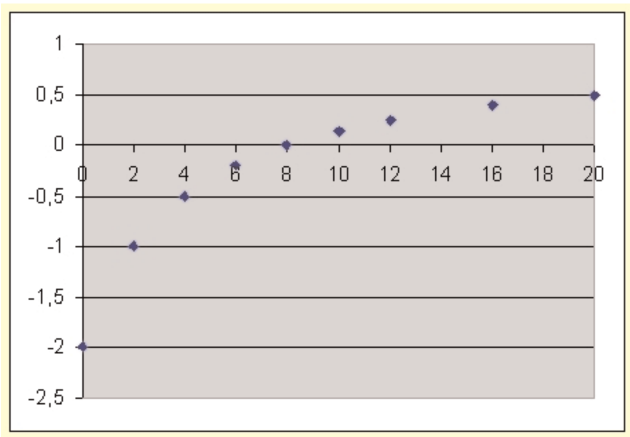
On note t le temps exprimé en mois et $g(t)$ le taux d'anticorps à l'instant t.



Déterminer graphiquement le sens de variation de la fonction g.
Que représentent les droites T et T' ?
Quelles sont les valeurs de $g'(4)$ et $g'(16)$?
Déterminer graphiquement $g'(8)$.
On peut tracer au jugé les tangentes aux points d'abscisses 0, 2, 6, 8, 10, 12 et 20, puis lire des valeurs approchées de leurs coefficients directeurs.

t	0	2	4	6	8	10	12	16	20
Coefficient directeur de la tangente	– 2	– 1	– 0,5	– 0,2	0	0,14	0,25	0,4	0,5

Coefficient directeur de la tangente



Quelle conjecture peut-on faire à propos du sens de variation de la fonction g' ?
Observer la position de la courbe par rapport à ses tangentes.
On admet que la dérivée de g est définie sur l'intervalle $[0\,;24]$ par $g'(t) = \dfrac{t-8}{t+4}$.

Retrouver le sens de variation de g.
En observant que $g'(t) = 1 - \dfrac{12}{t+4}$, établir le sens de variation de la fonction g'.

Le nombre $g'(t)$ représente la vitesse de l'évolution du taux d'anticorps dans le sang du nourrisson.

Des suites géométriques aux fonctions exponentielles

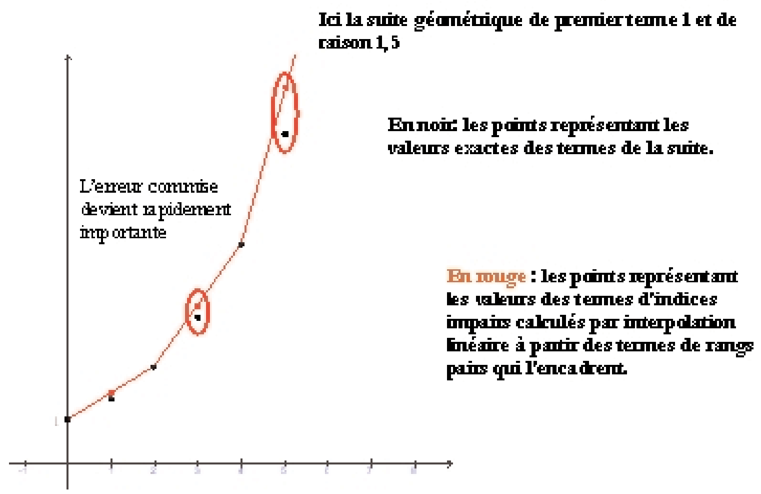
Commentaires sur le programme

La définition des suites arithmétiques et géométriques donnée dans le programme n'impose pas le mode d'introduction : il est tout à fait possible de donner une définition du type « Dire qu'une suite de nombres est arithmétique signifie qu'on passe d'un terme au suivant en ajoutant toujours le même nombre » puis de demander aux élèves comment traduire cette définition par une formule. Cette traduction n'a rien d'immédiat pour des élèves qui, découvrant la notation u_n, ont tendance à confondre u_{n+1} et $u_n + 1$. C'est l'une des raisons pour lesquelles le programme demande de citer la notation $u(n)$. Il y a deux autres raisons : c'est la notation utilisée par les calculatrices ; elle montre que les suites sont un cas particulier de fonctions.
Pour désigner une suite de nombres on disposera donc de deux notations : la notation (u_n) qui condense l'écriture $(u_o, u_1, u_2 \ldots)$ ou la notation u, plus simple mais plus abstraite.
Le premier objectif est de savoir reconnaître dans une situation donnée si l'on a affaire à une suite arithmétique, géométrique, ou d'une autre nature. En particulier les élèves doivent savoir que si la variation absolue est constante, il s'agit d'une suite arithmétique et que si la variation relative est constante, il s'agit d'une suite géométrique (dont la raison est non pas le taux, mais le coefficient multiplicateur).
Le passage de la formule d'itération à la formule explicite pose le problème du premier terme : est-il de rang 0 ou de rang 1 ?
Cela dépend de la situation étudiée : s'il s'agit de placer un capital à intérêts simples ou composés, il est commode d'appeler n la durée du placement, et donc C_o le capital initial ; mais si on effectue des versements à date fixe à une compagnie d'assurance, il est naturel d'appeler n le numéro du versement, et donc de noter le premier v_1.
Aussi, il peut être intéressant de montrer aux élèves les relations $u_n = u_p + (n - p)r$ et $u_n = u_p q^{n-p}$ où $n - p$ représente la différence des indices, afin que les élèves ne soient pas perturbés par cette question de relation démarrant à $n = 0$ ou $n = 1$.
L'utilisation de la calculatrice ou du tableur est ici très utile : il est en particulier formateur de comparer sur tableur la formule d'itération et la formule explicite.
La représentation graphique des suites pose la question : faut-il relier les points ? Cette question peut être soumise aux élèves et le professeur leur fera découvrir que la réponse est non pour deux raisons :
– les nombres réels autres que les entiers naturels n'ont pas d'image ;
– il y a une infinité de fonctions interpolant une suite donnée.
Ce problème de l'interpolation est d'ailleurs une question importante dans les sciences appliquées : elle sera abordée modestement en terminale avec l'ajustement affine.
On peut relier les points à l'axe des abscisses pour réaliser un diagramme en bâtons, mais il doit être clair qu'il s'agit simplement de faciliter la lecture (les points du « bâton » n'ont pas de signification).
Les mots « interpolation » et « extrapolation » ne sont pas au programme. On peut parler de courbe passant par des points donnés.
Pour une suite arithmétique il est intéressant de constater et de faire démontrer que les points sont alignés ; autrement dit, qu'il existe une fonction affine qui interpole la suite.
Pour une suite géométrique (a_n), ce n'est bien entendu pas le cas : on verra en terminale la fonction exponentielle $x \mapsto a^x$, qui permet l'interpolation la plus simple.

On veillera à établir le lien entre la suite $n \mapsto a^n$ et la fonction $x \mapsto a^x$. Les fonctions exponentielles fournissent un modèle continu là où les suites géométriques fournissent un modèle discret. Cela revient à chercher les fonctions qui transforment les sommes en produits, et donc une suite arithmétique en suite géométrique. On peut alors entre a^n et a^{n+1} intercaler leur moyenne géométrique, qui doit être l'image de $(n + 0{,}5)$. En itérant le procédé, on peut introduire a^x pour des valeurs non entières de x. On parle indifféremment de croissance (décroissance) exponentielle ou de croissance (décroissance) géométrique.

Pistes de réflexion pour le professeur

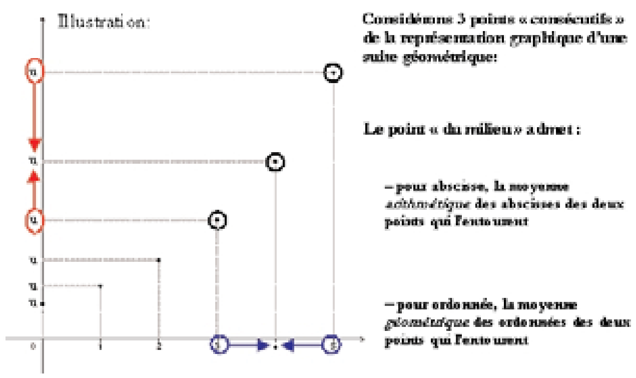
Vers les fonctions exponentielles



Construction d'une fonction exponentielle

Les fonctions exponentielles sont présentées comme le *prolongement* des suites géométriques de premier terme 1 et de raison q strictement positive.
La démarche est *expérimentale*.
Elle consiste à compléter le nuage de points représentant les puissances entières d'un réel strictement positif q.
L'algorithme de construction des points est basé sur le *principe de dichotomie*.
Il s'appuie sur les deux résultats suivants :
– Trois réels a, b et c sont, dans cet ordre, trois termes consécutifs d'une suite arithmétique si et seulement si b est la *moyenne arithmétique* de a et de c, c'est-à-dire $b = \frac{a + c}{2}$.

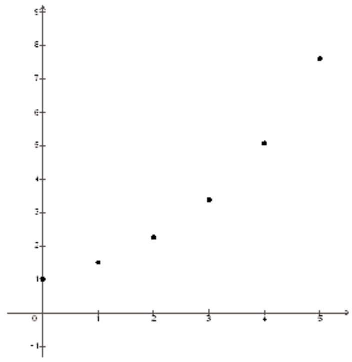
– Trois réels positifs a, b et c sont, dans cet ordre, trois termes consécutifs d'une suite géométrique si et seulement si b est la *moyenne géométrique* de a et de c, c'est-à-dire $b = \sqrt{ac}$.



Exemple : construction de la fonction $x \mapsto 1{,}5^x$ à partir de la suite géométrique de premier terme 1 et de raison 1,5.
Outils : tableur et grapheur.

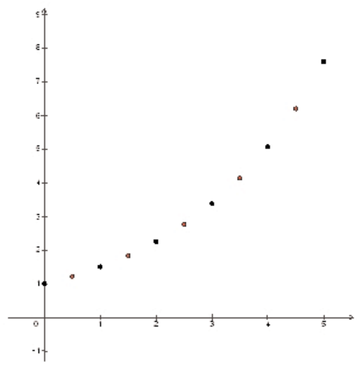
– 1re étape : points à abscisses entière

n	(1,5)^n
0	1
1	1,5
2	2,25
3	3,38
4	5,06
5	7,59
6	11,39
7	17,09
8	25,63
9	38,44



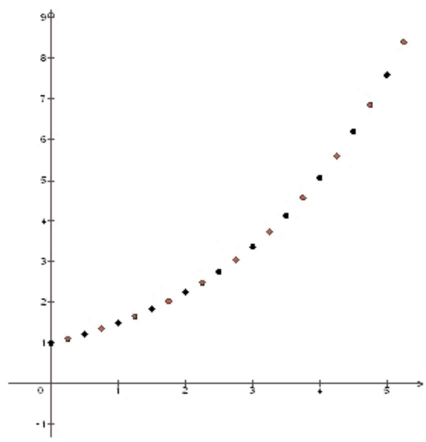
– 2e étape : points dont les abscisses sont de la forme $\left(n + \frac{1}{2}\right)$

n	(1,5)^n	x	y
0	1	0,5	1,22
1	1,5	1,5	1,84
2	2,25	2,5	2,76
3	3,38	3,5	4,14
4	5,06	4,5	6,2
5	7,59	5,5	9,3
6	11,39	6,5	13,95
7	17,09	7,5	20,93
8	25,63	8,5	31,39
9	38,44		



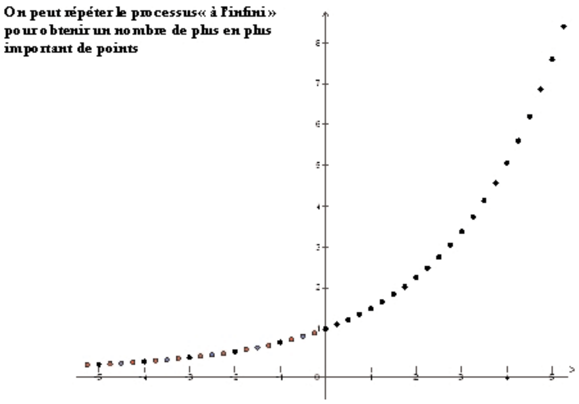
– 3e étape : points dont les abscisses sont de la forme $\left(n + \frac{1}{4}\right)$ et $\left(n + \frac{3}{4}\right)$

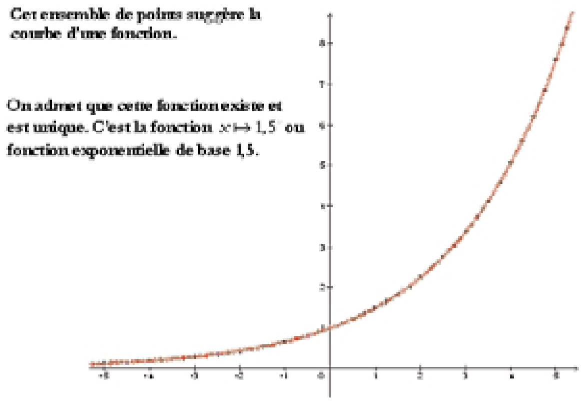
n	(1,5)^n	x	y
0	1	0,25	1,11
0,5	1,22	0,75	1,36
1	1,5	1,25	1,66
1,5	1,84	1,75	2,03
2	2,25	2,25	2,49
2,5	2,76	2,75	3,05
3	3,38	3,25	3,74
3,5	4,14	3,75	4,57
4	5,06	4,25	5,6
4,5	6,2	4,75	6,86
5	7,59	5,25	8,4
5,5	9,3	5,75	10,29
6	11,39	6,25	12,61
6,5	13,95	6,75	15,44
7	17,09	7,25	18,91
7,5	20,93	7,75	23,16
8	25,63	8,25	28,36
8,5	31,39	8,75	34,74
9	38,44		



Sachant que $1{,}5^{-n} = \left(\frac{1}{1{,}5}\right)^n$, on peut compléter le graphique en partant de la suite géométrique de premier terme 1 et de raison $\frac{1}{1{,}5}$.

On utilise le même processus dichotomique pour obtenir un nombre croissant de points.





Propriétés des fonctions exponentielles

Pour tout réel q strictement positif, la fonction exponentielle de base q est la fonction $x \mapsto q^x$.
Les propriétés suivantes sont admises :
Les fonctions $x \mapsto q^x$ sont définies et dérivables sur $\mathbb{R}$.
Pour tout réel x, q^x est strictement positif.
Pour tous réels x et y, $q^x \times q^y = q^{x+y}$.
Pour tout réel x, $q^{-x} = \dfrac{1}{q^x}$.

N.B. – L'expression de la dérivée de la fonction $x \mapsto q^x$, l'allure des courbes des fonctions exponentielles ainsi que leur comportement à l'infini ne font pas partie des objectifs du programme.

Pistes de réflexion

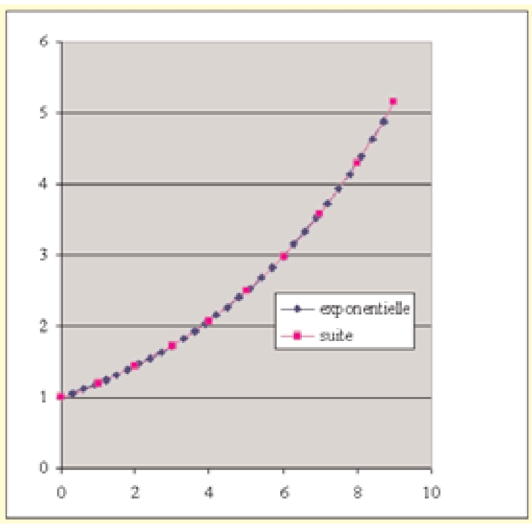
1. On a placé le 1er janvier 2005 la somme de 1 000 € sur un livret rapportant 3,5 % d'intérêts (composés) par an. De quelle somme pourra-t-on disposer le 1er mars 2008 ?
2. Valeur acquise par un capital C_0 placé à intérêts composés.

A	B	C	D	E	F	G	H	I		
	i : Taux d'intérêt annuel *n* : Durée, en années, du placement $C_0 =$ 2500,00									
i / *n*	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
1	2525,00	2550,00	2575,00	2600,00	2625,00	2650,00	2675,00	2700,00	2725,00	2750,00
2	2550,25	2601,00	2652,25	2704,00	2756,25	2809,00	2862,25	2916,00	2970,25	3025,00
3	2575,75	2653,02	2731,82	2812,16	2894,06	2977,54	3062,61	3149,28	3237,57	3327,50
4	2601,51	2706,08	2813,77	2924,65	3038,77	3156,19	3276,99	3401,22	3528,95	3660,25
5	2627,53	2760,20	2898,19	3041,63	3190,70	3345,56	3506,38	3673,32	3846,56	4026,28
6	2653,80	2815,41	2985,13	3163,30	3350,24	3546,30	3751,83	3967,19	4192,75	4428,90
7	2680,34	2871,71	3074,68	3289,83	3517,75	3759,08	4014,45	4284,56	4570,10	4871,79
8	2707,14	2929,15	3166,93	3421,42	3693,64	3984,62	4295,47	4627,33	4981,41	5358,97
9	2734,21	2987,73	3261,93	3558,28	3878,32	4223,70	4596,15	4997,51	5429,73	5894,87
10	2761,56	3047,49	3359,79	3700,61	4072,24	4477,12	4917,88	5397,31	5918,41	6484,36
11	2789,17	3108,44	3460,58	3848,64	4275,85	4745,75	5262,13	5829,10	6451,07	7132,79
12	2817,06	3170,60	3564,40	4002,58	4489,64	5030,49	5630,48	6295,43	7031,66	7846,07
13	2845,23	3234,02	3671,33	4162,68	4714,12	5332,32	6024,61	6799,06	7664,51	8630,68
14	2873,69	3298,70	3781,47	4329,19	4949,83	5652,26	6446,34	7342,98	8354,32	9493,75
15	2902,42	3364,67	3894,92	4502,36	5197,32	5991,40	6897,58	7930,42	9106,21	10443,12

3. Fonction $x \mapsto a^x$ et suite géométrique $n \mapsto a^n$.
Cas a = 1,2 avec un pas de 0,3 et un premier terme correspondant à x = 0.

x	a^x
0	1
0,3	1,05621997
0,6	1,11560062
0,9	1,17831965
1,2	1,24456475
1,5	1,31453414
1,8	1,38843721
2,1	1,4664951
2,4	1,54894141
2,7	1,63602285
3	1,728
3,3	1,82514811
3,6	1,92775787
3,9	2,03613636
4,2	2,15060788
4,5	2,27151499
4,8	2,39921949
5,1	2,53410354
5,4	2,67657076
5,7	2,82704748
6	2,985984
6,3	3,15385593
6,6	3,33116561
6,9	3,51844363
7,2	3,71625042
7,5	3,9251779
7,8	4,14585128
8,1	4,37893091
8,4	4,62511427
8,7	4,88513805

n	a^n
0	1
1	1,2
2	1,44
3	1,728
4	2,0736
5	2,48832
6	2,985984
7	3,5831808
8	4,29981696
9	5,15978035
10	6,19173642
11	7,43008371



Fonction logarithme décimal

Commentaires du programme

La fonction logarithme décimal est la seule exigible en ST2S, elle doit donner lieu à des activités variées dans un cadre pluridisciplinaire.
La première table de logarithmes décimaux a été proposée en 1617 par Briggs, disciple de Neper, comme une amélioration de la découverte de son maître. Plus commode pour les calculs numériques, les logarithmes décimaux ont été extrêmement utilisés jusqu'à l'apparition des calculatrices à la fin des années 1960.
Il en subsiste quelques usages en physique : pH d'une solution, magnitude d'une étoile, mesure en savarts de la hauteur d'un son, mesure en décibels de l'intensité d'un son… (voir la partie « Pistes de réflexion pour le professeur » ci-dessous).

Pistes de réflexion pour le professeur

Historique

La mise en relation d'une suite de puissances d'un nombre avec la suite correspondante des exposants, fondement de la théorie des logarithmes, remonte à l'époque paléo-babylonienne (XVIIIe siècle avant J.-C.) et non, comme le prétendaient certains historiens, à l'*Arénaire* d'Archimède. La contribution babylonienne a même été, par certains aspects, plus riche que celle d'Archimède, car elle considérait les puissances successives de différents nombres. On a retrouvé une tablette didactique, comprenant les puissances de 225, de 2 à 7 et où il est demandé de compléter jusqu'à 10. Une autre tablette reprend les puissances de 9. On possède aussi des tablettes pour les puissances de 16 et de 100. Une autre tablette répond à la question : à quelle puissance faut-il élever un certain nombre a pour obtenir un nombre donné ?
$16^{1:4} = 2$; $16^{1:2} = 4$; $16^{3:4} = 8$; $16^{1} = 16$; $16^{5:4} = 32$; $16^{3:2} = 64$.
Le point de départ de tout le problème traité par les Babyloniens doit probablement être trouvé dans les calculs d'intérêts. Les Babyloniens n'hésitaient pas à faire une interpolation linéaire pour trouver des valeurs approximatives… On trouve un exemple clair de l'usage pratique d'interpolation avec des tables exponentielles dans un texte – problème qui demande combien de temps il faut pour qu'un capital double si le taux d'intérêt annuel est de 20 %. La réponse est 3,47,13,20 soit 3 ans 47/60 13/3600 20/216000. Il semble très clair que le scribe a utilisé une interpolation linéaire entre les valeurs $(1\,;\,12)^3$ et $(1\,;\,12)^4$, d'après la formule d'intérêt composé $C(1 + r)^n$ où r est 20 %, c'est-à-dire 12/60.

On a retrouvé aussi un texte de Mari (XVIIIe siècle avant J.-C.) qui met en correspondance la suite des puissances de 2 et celle des exposants :
« 1 grain a fait augmenter 1 grain, soit 2 grains le 1er jour,
4 grains le 2e jour,
8 grains le 3e jour… »
et cela continue ainsi, mais les dernières lignes (28, 29, 30) sont fausses, peut-être à cause d'une mauvaise interprétation du résultat du calcul écrit en sexagésimal par le scribe (Imhotep).

En voulant compter le nombre de grains de sable qui se trouveraient contenus dans une sphère de la grandeur de notre univers, il trouve que c'est plus petit que 1 000 unités du 7e ordre de nombres, soit pour nous 10^{51}. Archimède (287-212 avant J.-C.) propose de représenter les grands nombres par des puissances de 10. Il n'est pas loin des logarithmes avec son astuce de calcul du produit des grandes puissances.
Il énonce la règle $a^m \times a^n = a^{m+n}$, sous une forme un peu différente, que nous pourrions traduire ainsi, dans nos notations modernes : « Dans la suite des nombres proportionnels 1, a^1, a^2, a^3…, a^{n-1}, a^n…, a^m, a^{m+1}…, a^{m+n}… où le rang de chaque nombre est égal à son exposant augmenté de 1, la distance du produit $a^m \times a^n = a^{m+n}$ à a^m est mesurée par $(n + 1)$ nombres et sa distance à l'unité par $(m + n)$ nombres. »

La méthode d'Archimède

On met en correspondance les nombres (progression **A**rithmétique) et les puissances de 2 (progression **G**éométrique) :

A	1	2	3	4	5	6	7	8	9	10
G	2	4	8	16	32	64	128	256	512	1024

Pour multiplier deux nombres de la ligne **G**, on ajoute les deux nombres parents sur la ligne **A**. On cherche le fiston correspondant à la somme trouvée.
Exemple :
A $3 + 4 = 7$
G $8 \times 16 = 128$
Aujourd'hui, on sait en effet, que : $2^x \times 2^y = 2^{x+y}$.
Avec les puissances de 10 :

N	100	200	300	400	500	600	700	800	900	1000
$log_{10}N$	2	2,3	2,5	2,6	2,7	2,78	2,85	2,9	2,95	3
Différence		0,3	0,2	0,1	0,1	0,08	0,07	0,05	0,05	0,05

La dernière ligne montre le tassement des valeurs vers la grande valeur en échelle logarithmique.
L'astronome Kepler (1571-1630) avait tenté d'alléger les calculs.

Les « os de Napier »

Les « os de Napier » sont des bâtons sur lesquels étaient gravées les tables de multiplication qu'on pouvait disposer en un réseau permettant de lire rapidement le résultat d'une multiplication, même avec des grands nombres. C'est un principe qui transforme les multiplications en additions. L'idée des logarithmes est née de cette même idée : accélérer les calculs.

Premières observations

Considérons un nombre positif a et observons avec les notations modernes :

a	a^2	a^3	a^4	a^5	…	a^n	…
1	2	3	4	5	…	n	…

a) Quelle est la nature des deux suites construites ?
b) Nicolas Chuquet dans son *Triparty en la science des nombres* (1484), remarqua que si l'on fait correspondre deux termes de même rang de chacune des deux suites au produit de deux termes de la première est associé la somme des termes correspondants de la seconde. Prouver ce résultat.
c) Que peut-on dire des termes de la seconde suite ?
C'est en généralisant la remarque de Nicolas Chuquet que Stiffel (1486-1567) dans son traité *Arithmetica Integra* introduisit, en 1544, les exposants négatifs. Il n'hésita pas à utiliser les nombres négatifs et alla jusqu'à écrire les progressions arithmétiques et géométriques :

… – 5	-4	-3	-2	-1	0	1	2	…
… 1/32	1/16	1/8	1/4	1/2	1	2	4	…

John Napier (ou Neper)

Écossais, né en 1550, Neper était un protestant convaincu à une époque où catholiques et protestants se livraient à une guerre civile larvée. Il publia en 1593 un ouvrage sur l'Apocalypse qui eut un grand succès.
Il proposa un système de réglettes facilitant les multiplications.
Suivant la même démarche de simplification des calculs, il travailla pendat vingt ans à l'élaboration d'un nouvel outil. Essayons, en partant des progressions de Stiffel, de reconstituer la démarche de Neper qui lui demanda vingt années d'efforts.

Soit a un nombre quelconque strictement supérieur à 1.
Écrivons les deux suites, infinies dans les deux sens :

…	a^{-3}	a^{-2}	a^{-1}	1	a	a^2	a^3	a^4	a^5	…	a^n	…
…	–3	–2	–1	0	1	2	3	4	5	…	n	…

On dira que a est la base, c'est-à-dire ici la raison de la suite géométrique.
À tout terme de la première suite, on peut associer un unique terme de la seconde et réciproquement.
Neper invente alors le mot « logarithme » du grec *logos* : logique, raison et *arithmos* : nombre.
Neper dit ainsi que 0 est le logarithme de 1, – 1 est le logarithme de a^{-1}…
La notation utilisée fut : *L.x* ; *l.x* ; *Log.x* ; *L :x* ; *cd logx* ; */x* ; *LPx* ; ^{a}*log.x*
1. Combien de logarithmes existe-t-il ? De quoi dépendent-ils ?
2. Soit a un réel strictement positif. Quelle est la valeur de son logarithme en base a ?

Exemple d'utilisation d'une table de logarithmes.

Nombre	*log*	*Nombre*	*log*	*Nombre*	*log*
0,000 064	–6	5,000 000	1	390 625	8
0,000 320	–5	25,000 000	2	1 953 125	9
0,001 600	–4	125,000 000	3	9 765 625	10
0,008 000	–3	625,000 000	4	48 828 125	11
0,040 000	–2	3 125, 000 000	5	244 140 625	12
0,200 000	–1	15 625, 000 000	6	1 220 703 125	13
1,000 000	0	78 125, 000 000	7	6 103 515 625	14

1. Dans quelle base est complétée la table de logarithmes ci-dessus ?
2. Effectuer les opérations suivantes en utilisant la table :
$0{,}04 \times 3\,125$; $3\,125 \times 1\,953\,125$; $0{,}000\,064 \times 390\,625$; $0{,}04 \times 0{,}001\,6$; $78\,125 \times 625$.
3. Citer les avantages et les inconvénients de cette méthode.

Compléter une table de logarithmes

Nous allons maintenant étudier un texte du mathématicien français Ozanam (1640-1717) qui nous permettra de comprendre comment compléter une table de logarithmes.
Préambule :
Soient a et b deux réels non nuls de logarithmes respectifs α et β. Le problème est de déterminer en fonction de a et b un réel c tel que c ait pour logarithme $\frac{\alpha + \beta}{2}$. c est appelé moyenne géométrique de a et b.

Texte d'Ozanam

« Les logarithmes sont des nombres en proportion arithmétique, correspondant à d'autres nombres en proportion géométriques, desquels ils sont appelés logarithmes. Comme il est libre de prendre telle progression que l'on voudra, on choisira la plus commode, qui est de prendre la progression décimale pour la progression géométrique et la progression des nombres naturels pour l'arithmétique, en sorte que, pourtant, le premier nombre arithmétique, qui répond au premier géométrique, ou à l'unité, soit 0, c'est-à-dire que le logarithme de l'unité soit Ø, pour rendre l'usage des logarithmes plus facile : comme vous le voyez dans cette table, où le logarithme de 1 est 0, de 10 est 1, de 100 est 2, de 1000 est 3 et ainsi de suite ; et parce que, dans la pratique, on a besoin des logarithmes des nombres moyens 2, 3, 4, 5, etc., et que ces logarithmes ne peuvent être exprimés qu'en fractions, on se servira aussi de la progression décimale pour la facilité du calcul, en ajoutant un certain nombre de zéros à chaque terme de la progression arithmétique, plus ou moins exacts, comme vous voyez ici.

Ainsi, nous supposerons que le logarithme de 10 est 1,000000, que le logarithme de 100 est 2,0000000, celui de 1 000 est 3,0000000, etc., en suite de quoi il faut trouver les logarithmes des nombres moyens 2, 3, 4, 5, etc., ce que nous ferons après avoir expliqué la nature et les propriétés des logarithmes dans les propositions suivantes :

Proposition.

La somme des logarithmes de deux nombres entiers est égale au logarithme de leur produit, lorsque le logarithme de l'unité est 0.

Proposons par exemple les deux nombres entiers 4, 6, dont le produit est 24. Je dis que le logarithme de 24 est égal à la somme des logarithmes de 4 et de 6, le logarithme de l'unité étant 0. Car, puisque 24 est le produit de 4 et de 6, ces quatre nombres 1, 4, 6, 24 seront en proportion géométrique, c'est pourquoi leurs logarithmes seront en proportion arithmétique, et la somme des deux extrêmes, c'est-à-dire la somme des logarithmes de 1 et de 24, sera égale à la somme des deux moyens ou à la somme des logarithmes de 4 et de 6, et parce qu'on suppose que le logarithme de 1 est 0, le seul logarithme de 24 sera égal à la somme des logarithmes de 4 et de 6, qui produisent 24. Ce qu'il fallait démontrer. […]

Pour trouver le logarithme d'un nombre donné, comme par exemple de 9 : parce que ce nombre 9 est entre les nombres 1 et 10, dont les logarithmes sont connus, savoir 0,0000000 et 1,0000000, ou 0,00000000, 1,00000000, en les augmentant chacun d'un zéro, pour avoir plus exactement le logarithme qu'on cherche, faites ainsi.

Entre 1 et 10, augmentez d'autant de zéros que leurs logarithmes en contiennent, comme ici de sept zéros, pour avoir exactement dans le même nombre de figures le logarithme qu'on cherche. Savoir entre A et B, trouvez un moyen géométrique C lequel étant moindre que 9,0000000, il faudra chercher entre le moindre C et le plus grand B, un autre moyen proportionnel D, qui est encore moindre que 9,0000000 : c'est pourquoi, entre le moindre D et le plus grand B, on cherchera un troisième moyen proportionnel E, entre lequel et le plus grand B, on trouvera un quatrième moyen proportionnel F, qui est ici encore moindre que 9,0000000 : c'est pourquoi il faudra trouver, entre le moindre F et le plus grand B, un cinquième moyen proportionnel G, qui se rencontre ici plus grand que 9,0000000, ainsi entre le plus grand G et le prochainement moindre F, on cherchera un sixième moyen proportionnel H qui, étant moindre que 9,0000000, on doit chercher entre ce moindre H et le prochainement plus grand G, un septième moyen proportionnel I, qui est bien plus grand que 9,0000000, mais non pas avec un si grand excès comme le précédent G. C'est pourquoi, en cherchant entre le prochainement moindre et le prochainement plus grand des moyens géométriques proportionnels, on aura des nombres qui approchent toujours de plus en plus du nombre proposé 9,0000000, lequel enfin se rencontre ici le vingt-sixième moyen proportionnel : après quoi il sera facile de reconnaître son logarithme, car, comme entre les nombres A, B, nous avons trouvé un moyen géométrique proportionnel C, si entre leurs logarithmes on trouve un moyen arithmétique proportionnel, celui-ci sera le logarithme du premier moyen géométrique proportionnel C. C'est de la même manière que l'on trouvera les logarithmes de tous les autres moyens géométriques proportionnels et, par conséquent, du dernier 9,0000000, ou du nombre proposé 9, dont le logarithme se trouvera tel, 0,95424251.

C'est de cette manière qu'on trouvera les logarithmes des autres nombres entre 1 et 9, et des nombres entre 1 et 100, et des nombres entre 100 et 1000, et ainsi de suite : mais on pourra trouver avec beaucoup de facilité les logarithmes des nombres composés, par l'addition des logarithmes des nombres qui les composent, comme il est évident par ce qui a été dit dans la proposition. »

Pistes de réflexion pour le professeur

Le pH d'une solution (voir également page 64)

L'acidité d'une solution est mesurée par son pH, défini en fonction de la concentration en ions H_3O^+ par :
$pH = -\log[H_3O^+]$ où $[H_3O^+]$ est la concentration (en mole par litre) d'ions H_3O^+ de la solution.
Le pH de l'eau pure est 7 ; celui d'une *solution acide* est strictement inférieur à 7 alors que celui d'une *solution basique* est strictement supérieur à 7.
Le pH d'un sol, généralement compris entre 3,5 et 9,5 renseigne sur le degré de saturation du complexe absorbant. La fougère, le châtaignier recherchent un sol au pH voisin de 4 à 5 ; le buis exige un sol au pH de l'ordre de 8.

1. Quel est le pH d'une solution dont la concentration en $[H_3O^+]$ est $4{,}8 \times 10^{-6}$ mole par litre ?
Si la concentration est égale à $4{,}8 \times 10^{-6}$ mol.L^{-1}, alors le pH est donné par : $pH = -\log(4{,}8 \times 10^{-6})$ or $\log(4{,}8 \times 10^{-6}) = \log(4{,}8) + \log(10^{-6})$ soit environ $0{,}681 - 6$ soit environ $-5{,}319$. Le pH est donc environ 5,319.
2. Si le pH d'une solution est 4, quelle est la concentration en ions $[H_3O^+]$?
Si le pH est 4 alors $4 = -\log[H_3O^+]$ d'où $\log[H_3O^+] = -4 = \log(10^{-4})$. Ainsi $[H_3O^+] = 10^{-4}$ mole par litre.
3. Que fait le pH lorsque la concentration est divisée par 10 ?
Posons $pH = -\log[H_3O^+]$ et $pH' = -\log\left(\frac{[H_3O^+]}{10}\right)$. Alors $pH' = -\log[H_3O^+] + \log(10)$, soit pH + 1. Si la concentration est divisée par 10, le pH augmente de 1.

L'intensité sonore

Pour une fréquence donnée, l'oreille humaine est capable de percevoir un son dont on mesure le niveau d'intensité S en décibel (dB) par $S = 10\log\left(\frac{I}{I_0}\right)$ où I est l'intensité sonore du son étudié alors que I_0 est l'intensité sonore d'un son de référence (fixé par l'American Standart Association).

La pression acoustique

Le niveau de pression acoustique d'un son, NPA, mesuré en décibels (dB), est fonction de la pression acoustique P mesurée en pascals (Pa).
NPA = $f(P)$ avec $f(P) = 20\log\left(\frac{P}{2.10^{-5}}\right)$.

Pour une fréquence de 1 000 Hertz :
– le seuil de perception d'un son (seuil en dessous duquel le son est inaudible) correspond à une pression acoustique P égale à $3{,}2 \times 10^{-5}$ Pa ;
– le seuil de la douleur (seuil maximum supportable) correspond à une pression acoustique P égale à 64 Pa.

De la statistique aux probabilités

Commentaires du programme

La statistique

La statistique est l'art de caractériser une population à partir de mesures sur les individus qui la composent. Le choix de la caractéristique dépend de l'objectif poursuivi (voir utilisation de la médiane et de la moyenne).
D'une part, elle peut être reliée à la partie « Information chiffrée » par son aspect descriptif ; d'autre part, elle s'appuie sur la théorie des probabilités pour le choix et la validation d'un modèle.
On peut travailler sur des données réelles que les élèves peuvent recueillir par enquête dans leur lycée, dans une entreprise, une administration ou par investigation dans la presse ou sur Internet et qui ont un sens pour eux.
L'utilisation de la calculatrice ou du tableur permet le traitement d'un nombre important de données, nécessaire pour que l'étude statistique se justifie ; c'est pourquoi il faut consacrer un temps suffisant pour que les élèves maîtrisent, avec ces deux outils, les fonctions statistiques à leur programme. Une des difficultés est l'introduction des données :
– pour une variable, il faut savoir indiquer où sont les valeurs de la variable et où sont les effectifs ;
– pour deux variables, il faut savoir indiquer où sont les valeurs de chaque variable.
Les nouveautés de la classe de première concernent, d'une part, les tableaux de contingence (voir « Information chiffrée »), d'autre part les paramètres de dispersion (écart type, écart interquartile).
D'une façon générale, le sens de la démarche est à privilégier. C'est pourquoi le programme insiste sur la rédaction de l'interprétation du résumé statistique obtenu, qu'il s'agisse du couple (moyenne ; écart type) ou du couple (médiane ; écart interquartile) : la moyenne et la médiane mesurent ce qu'il est convenu d'appeler la *tendance centrale* ou la *position* de la série statistique, alors que l'écart type ou l'écart interquartile en mesurent la *dispersion*. C'est souvent par la comparaison de deux ou plusieurs séries que ces notions prennent du sens : qu'on pense par exemple à deux villes de même température moyenne, l'une au climat continental et l'autre au climat océanique ; ou à deux classes de première ST2S de même niveau médian, l'une homogène et l'autre hétérogène ; ou encore à la distribution des fréquences de succès dans deux échantillons d'une même épreuve répétée. Une remarque concernant les regroupements de données en classes : le professeur doit veiller à ce que les élèves comprennent qu'à partir de données regroupées en classes, on ne peut obtenir que des approximations des paramètres statistiques puisque l'on a perdu de « l'information ». Les moyens techniques actuels permettant de traiter un très grand nombre de données, il convient de privilégier le traitement de données réelles discrètes pour déterminer ces paramètres. Cependant, le regroupement de données en classes peut être pertinent pour tracer une représentation graphique qui donne un sens à la distribution de la population étudiée.
En ce qui concerne les quantiles, le professeur lira avec profit le document d'accompagnement des programmes de la classe de première des séries générales. La définition des quantiles est donnée dans le lexique de ce document. Cette définition est difficilement accessible aux élèves. On peut d'ailleurs noter que de nombreux statisticiens,

de nombreux logiciels préfèrent utiliser un algorithme pratique. Par exemple, pour trouver la médiane : on ordonne la série des observations par ordre croissant ; si la série est de taille $2n + 1$, la médiane est la valeur du terme de rang $n + 1$ dans cette série ordonnée ; si la série est de taille $2n$, la médiane est la demi-somme des valeurs des termes de rang n et $n + 1$ dans cette série ordonnée. On peut reprendre la même démarche pour trouver les quartiles. Cette dernière définition de la médiane est celle adoptée dans le programme de seconde. Elle ne fournit pas toujours exactement le même résultat que la définition officielle ou que certaines calculatrices ou certains tableurs. Mais ces deux définitions donnent en pratique, pour des séries de grande taille (et ces notions ont un intérêt précisément dans ce contexte), des résultats très proches : leurs différences ne présentent pas d'intérêt au vu de la définition de la statistique fournie auparavant.

Il est préférable de distinguer l'*intervalle* interquartile $[Q_1 ; Q_3]$ de l'*écart* interquartile $Q_3 - Q_1$: utiliser le même mot ne pose aucun problème au statisticien chevronné, mais peut être source d'erreur pour l'élève de première. En outre, le mot écart souligne l'analogie avec l'écart type.
La médiane et les quartiles ont l'avantage de n'être pas sensibles aux valeurs extrêmes, contrairement à la moyenne et l'écart type : il faut faire comprendre aux élèves que, selon le contexte de l'exercice, il est préférable d'utiliser le couple moyenne/écart type ou le couple médiane/écart interquartile.
L'étude des revenus d'une population d'un pays en est un bon exemple. Si on est en économie, moyenne et écart type sont les indicateurs adaptés car ils sont liés à la richesse moyenne de la nation ; si on est en sociologie, l'individu percevant le revenu moyen sera vu comme un riche car il est, en général, nettement au-dessus du revenu médian. Dans ce dernier cas, le couple d'indicateurs médiane/écart interquartile ou interdécile est plus pertinent.
Actuellement les outils informatiques fournissent automatiquement la moyenne et l'écart type ; les élèves doivent savoir interroger leur calculatrice pour les obtenir : cela appelle un apprentissage spécifique et cela ne dispense pas de faire un ou deux calculs à la main pour en comprendre le principe.
Le *diagramme en boîte* (dit aussi « diagramme de Tuckey », « boîte à moustaches » ou « boîte à pattes ») peut s'arrêter aux valeurs extrêmes ou aux derniers déciles selon les attentes précisées dans l'énoncé. Il permet des comparaisons visuelles rapides entre plusieurs séries statistiques, tant en ce qui concerne la tendance centrale que la dispersion (voir fiche 63 du site Euler).
On étudiera des données recueillies par les élèves, tout en choisissant des situations permettant de limiter le temps de recueil de ces données. Si les effectifs des données recueillies restent faibles, on aura toujours intérêt à se procurer les mêmes données pour un niveau régional ou national, *via* le site de l'INSEE par exemple.
À cette occasion, on s'attachera à :
– définir une problématique ou une question précise motivant un recueil de données expérimentales ;
– définir les données à recueillir, leur codage et les traitements statistiques qu'on appliquera pour avoir des éléments de réponses à la question posée ;
– élaborer un protocole de recueil et aborder les problèmes que cela pose.
Proposition d'exemples : battements cardiaques, estimation de longueurs, durée des repas du soir, nombre et durée de conversations téléphoniques, temps de passage en caisse dans une grande surface etc.

En classe terminale, sont introduites les séries de données statistiques à deux variables. L'objectif est d'étudier le lien éventuel entre deux caractères d'une même population. Pour deux variables qualitatives, on peut étudier leur dépendance à l'aide des fréquences conditionnelles (voir « Information chiffrée ») et établir un lien avec le programme de probabilités (voir le conditionnement).
Pour deux variables quantitatives, on peut rechercher une formule approchée exprimant l'une des variables en fonction de l'autre : c'est la problématique de l'*ajustement*.

En probabilité

Quand on peut identifier la variabilité de mesures au résultat du hasard, on peut alors introduire la probabilité.

Pour étudier une épreuve aléatoire, on a besoin d'un *modèle*, qui précise d'une part les issues (on les suppose ici en nombre fini), d'autre part la distribution de probabilité entre ces issues. Or, mis à part les jeux de hasard où des hypothèses de symétrie fournissent *a priori* les probabilités des issues, c'est l'observation statistique antérieure qui permet d'estimer ces probabilités. Cette observation consiste à répéter l'épreuve un grand nombre de fois, et à relever dans cet *échantillon* les fréquences d'apparition des issues. Comme les divers échantillons fournissent des résultats différents, il est nécessaire d'étudier la fluctuation d'échantillonnage, autrement dit la variabilité entre les divers échantillons possibles de même taille (voir les fiches 152 et 153 du site Euler).

Le résultat essentiel, conforme à l'intuition, peut s'énoncer de façon qualitative en deux phrases :

– la distribution des fréquences fluctue autour de la distribution de probabilité ;

– l'amplitude de la fluctuation est d'autant plus faible que la taille de l'échantillon est grande.

Ces deux affirmations sont des théorèmes de la théorie des probabilités. L'observation en permet une validation *a posteriori*. Cela sort largement du cadre du programme des classes de première et terminale ST2S. Néanmoins, le travail expérimental fait en classe de seconde sur la fluctuation d'échantillonnage a sensibilisé les élèves à cette approche.

En classe de première, une façon possible de présenter la notion de probabilité est de dire : « Tout se passe comme si on tirait, au hasard, une boule dans une urne contenant des boules indiscernables au moment du tirage mais de diverses couleurs, chaque couleur correspondant à une issue. »

L'expression « au hasard », comme le mot « indiscernables » dans la phrase précédente, signifie, par convention, que chaque boule a « autant de chances » d'être choisie qu'une autre. La proportion de chaque couleur, autrement dit la probabilité de réalisation de chaque issue, est soit donnée *a priori*, soit estimée au préalable par répétition de l'épreuve (comme indiqué ci-dessus). Cette présentation a plusieurs avantages :

– elle fournit une image mentale de la probabilité et une méthode qui s'avère très efficace pour résoudre les problèmes, surtout quand on l'enrichit en traitant les tirages multiples (successifs avec remise, successifs sans remise) ;

– elle rend naturel le fait que la probabilité d'un événement soit la somme des probabilités de ses issues ;

– elle permet de réinvestir directement les propriétés des proportions vues dans le chapitre « Information chiffrée » ;

– elle permet de s'appuyer sur le travail fait en seconde reliant la proportion de boules rouges dans l'urne et la fréquence d'apparition d'une boule rouge lors de tirages répétés ;

– elle fournit une méthode générale pour la simulation : si l'urne contient N boules, tirer une boule équivaut à choisir au hasard un entier entre 1 et N.

Les tirages (simples ou multiples) dans une urne apparaissent ainsi comme les épreuves de référence, jouant le même rôle en probabilités que les fonctions de référence en analyse, les configurations-clés ou les transformations usuelles en géométrie.

Pour les tirages successifs, avec ou sans remise, on se limite à deux ou trois répétitions, l'objectif du programme étant de travailler sur les probabilités et en aucun cas de traiter du dénombrement. L'essentiel est en effet de faire comprendre :

– que l'équiprobabilité des boules dans l'urne induit l'équiprobabilité des couples ou des triplets ;

– que, dans le cas de tirages avec remise, contrairement au tirage sans remise, on répète plusieurs fois la même épreuve (on dit en terminale que les tirages sont *indépendants*).

L'utilisation des arbres probabilistes permet aux élèves de visualiser la succession des différentes issues possibles. On se limite en première à des arbres à branches équiprobables.

Les tirages simultanés ne sont pas au programme du cycle terminal : les situations où ils se présentent peuvent être traitées par des tirages successifs sans remise.

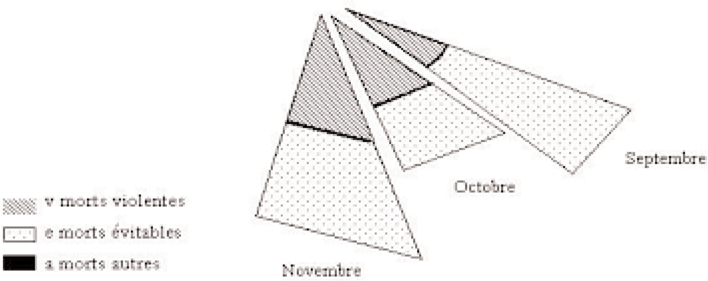
La notion de *probabilité conditionnelle* introduite en classe terminale prolonge naturellement celle de fréquence conditionnelle. Les événements A et B sont indépendants si la réalisation de B ne modifie pas la valeur de la probabilité de A d'un modèle à l'autre : $p_B(A) = p(A)$. Notons que la notation $p(A/B)$, usuelle en mathématiques appliquées, masque ce fait que p_B est une distribution de probabilité et peut faire croire à tort que « A/B » est un événement.
L'analyse de tableaux d'effectifs à deux entrées est un champ d'application important de la notion de probabilité conditionnelle. Un autre champ d'application est l'étude des *épreuves successives* : en les modélisant par des tirages successifs dans une urne, on est conduit à distinguer les tirages avec ou sans remise, ce qui permet de bien différencier l'indépendance et la dépendance.
Les arbres probabilistes sont un outil très efficace pour visualiser les propriétés et traiter les problèmes. En première, seuls des arbres à branches équiprobables, donc non pondérées, ont été utilisés. On peut faire remarquer que les arbres à branches pondérées sont une façon de résumer les arbres non pondérés par regroupement des branches. On peut aussi les rapprocher des arbres de répartition vus en première à l'occasion des tableaux de contingence.
Trois propriétés sont très utiles pour construire les arbres et les utiliser :
1. La probabilité d'un chemin est le produit des probabilités de ses branches.
2. Pour les branches issues d'un même nœud, la somme des probabilités vaut 1.
3. La probabilité d'un événement est la somme des probabilités des chemins qui y aboutissent.
En particulier, dans le cas d'un arbre à deux branches, ces propriétés se traduisent ainsi :
$p(A \cap B) = p_B(A) \times p(B)$; $p_B(A) + p_B(\overline{A}) = 1$; $p(A) = p(B) \times p_B(A) + p(\overline{B}) \times p_{\overline{B}}(A)$.
La première formule découle de la définition de la probabilité conditionnelle ; la deuxième du théorème des probabilités totales $p(A \cap B) + p(\overline{A} \cap B) = p(B)$; la troisième du même théorème, écrit sous la forme $p(A \cap B) + p(A \cap \overline{B}) = p(A)$.
On admet que ces propriétés restent vraies quel que soit le nombre de niveaux de l'arbre, et quel que soit le nombre de branches issues d'un nœud.
Lors d'évaluations, l'utilisation correcte d'un arbre bien construit doit être acceptée comme démonstration.
Un point délicat est la notion de dépendance. Le sens donné au mot dépendant dans le langage courant : « qui ne peut se réaliser sans l'action préalable d'une personne ou d'une chose » peut constituer une première approche. Il convient néanmoins de faire remarquer l'ambiguïté que peut induire cette « définition » dans le domaine mathématique : l'écriture « probabilité, sachant B, de A » fait dépendre de la réalisation de B la probabilité de la réalisation de A et non pas la réalisation de A elle-même. D'où la nécessité des formules caractérisant l'indépendance.
Le professeur trouvera une réflexion sur les liens entre dépendance causale et stochastique dans la partie « Statistique et probabilités » du document d'accompagnement des programmes des classes de terminale ES.
La formule de Bayes n'est pas exigible sous sa forme générale. Cela n'empêche pas de traiter quelques exemples de situations où, connaissant $P(B)$, $P_B(A)$ et $P_B(\overline{A})$, on cherche $P_A(B)$. Cela permet d'aborder la problématique de la « probabilité des causes » où, de façon qui peut sembler paradoxale, on s'intéresse à la probabilité d'un événement passé.

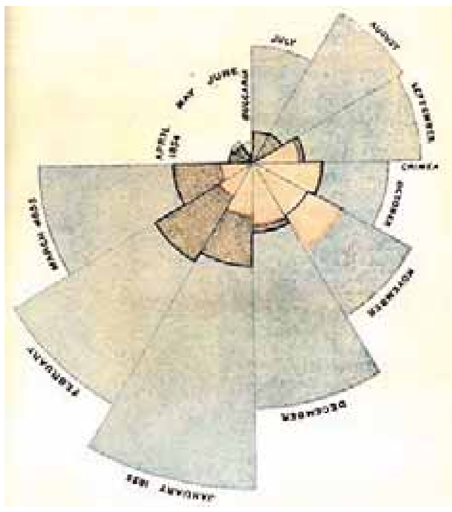
N.B.
– Un chemin est l'événement réalisé en suivant de gauche à droite des branches successives.
– Un nœud est la jonction de deux ou plusieurs branches.
– Sur ces questions de probabilités conditionnelles et d'indépendance, on lira avec profit l'ouvrage *Enseigner les probabilités au lycée*, commission Inter-IREM, Statistique et Probabilités, édité par l'IREM de REIMS. On y trouvera également des idées d'activités.

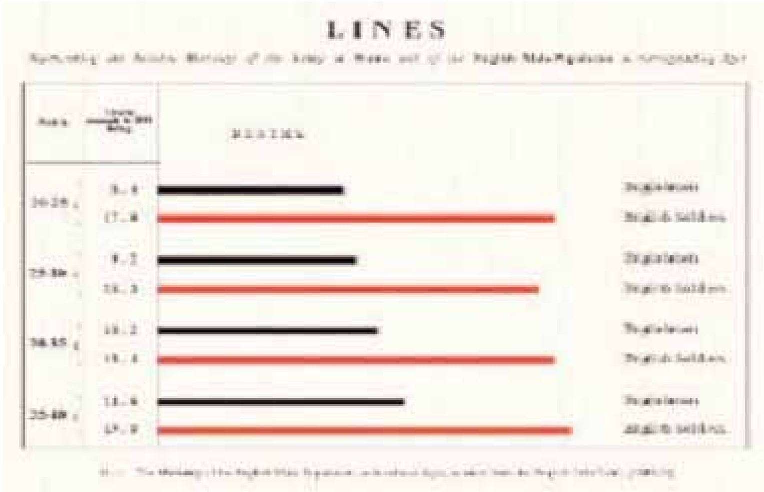
Pistes de réflexion pour le professeur

Un brin d'histoire…

Dans une section ou la santé joue un rôle central, Florence Nightingale (1820-1910) peut être considérée comme une référence. Florence Nightingale est remarquable en tant que pionnière de l'hygiène hospitalière, comme statisticienne, et comme pionnière de la formation des infirmières.
Fille d'une famille anglaise aisée, elle fait des études dans différents domaines (musique, langues, médecine, mathématiques…). Après avoir été surveillante générale dans un hôpital londonien elle part en 1854 sur le front de la guerre de Crimée avec un groupe d'infirmières volontaires. Touchée par la grande mortalité des soldats et persuadée que le manque d'hygiène constitue un facteur aggravant elle entreprend d'étudier les causes de leurs décès.
Elle comptabilise ainsi le nombre de morts en les classifiant en trois catégories : les morts violentes, les morts évitables, les autres morts. Elle présente ses résultats dans un diagramme en crêtes de coq (figures ci-dessous). De retour en Grande-Bretagne, Florence Nightingale compare la mortalité des hommes dans l'armée (qui n'était alors plus en guerre) avec celle des civils britanniques du même âge (page suivante). Ses résultats sont alors présentés sous forme d'un diagramme en bâtons.
Nouveaux à l'époque, ces graphiques basés sur le principe simple de la proportionnalité sont lisibles par tous. Ils démontrent que la plupart des soldats ne meurent pas des suites de leurs blessures mais de maladies qui auraient pu être évitées. Leur publication dans la presse provoque un scandale : l'état-major est contraint de réformer son organisation du système de santé.
Toute sa vie, Florence Nightingale s'attache à promouvoir une gestion hospitalière de qualité, des conditions de prise en charge des patients optimales et le développement des statistiques médicales. Elle s'inspire des travaux du statisticien Adolphe Quételet (1796-1874) et publie plusieurs livres et articles. Elle est la première femme à entrer au Statistical Institute.
Elle crée une école d'infirmières posant les bases d'une formation moderne.
Avec entre autre la prolifération des maladies nosocomiales, le combat de Florence Nightingale est toujours d'actualité. Elle continue d'occuper une place importante dans la communauté infirmière. La journée mondiale des infirmières a été fixée le 12 mai, jour de sa naissance.
Diagramme en crêtes de coq : Florence Nightingale représente chaque mois d'avril 1854 à mars 1855 par un secteur angulaire de mesure constante de 30 degrés ($12 \times 30 = 360$) Dans chaque secteur figure une aire proportionnelle à chaque type de décès (les morts violentes, les morts évitables, les autres morts) des soldats britanniques pendant la guerre de Crimée.







Caractéristiques de position et de dispersion : quelques précisions

Pour résumer une population on utilise souvent une valeur centrale associée à une mesure de dispersion : la moyenne associée à l'écart type ou la médiane associée à l'écart interquartile.
Les caractéristiques de centralité : médiane et moyenne ont été vues en seconde, celles de dispersion : écart type et écart interquartile sont des notions nouvelles en première.
Les définitions choisies sont celles employées dans les classes de première des sections générales ; on pourra se référer aux documents d'accompagnement de ces classes.

Définitions

Médiane

• Médiane d'un caractère discret : une fois les valeurs de la variable rangées par ordre croissant chacune figurant un nombre de fois égal à son effectif :
– si la série est de taille $N = 2n + 1$, une médiane est la *valeur* du terme de rang $n + 1$;
– si la série est de taille $N = 2n$, une médiane est la *moyenne entre les valeurs* des termes de rangs n et $n + 1$.

• Médiane d'un caractère continu : dans le cas particulier où on ne dispose que de données regroupées en classes, on peut déterminer, graphiquement ou par interpolation, une *valeur approchée* d'une médiane sur le polygone des fréquences cumulées croissantes (ou des effectifs cumulés croissants).

Quantiles : ayant rangé les termes de la série par ordre croissant :

• Premier quartile : c'est la plus petite valeur q des termes de cette série telle qu'au moins 25 % des données soient inférieures ou égales à q.

Par exemple, si la série est de taille 16 le quartile 1 est la quatrième valeur $\left(\frac{25}{100} \times 16\right) = 4$;

si la série est de taille 17 le quartile 1 est la cinquième valeur (car : $4 \leq \frac{25}{100} \times 17 \leq 5$).

• Troisième quartile : c'est la plus petite valeur q des termes de cette série telle qu'au moins 75 % des données soient inférieures ou égales à q.

• Premier décile : c'est la plus petite valeur q des termes de cette série telle qu'au moins 10 % des données soient inférieures ou égales à q.

• Neuvième décile : c'est la plus petite valeur q des termes de cette série telle qu'au moins 90 % des données soient inférieures ou égales à q.

De même pour les autres déciles.
On peut aussi définir les centiles.
Toutes ces définitions proviennent d'une même notion : la fonction quantile : Q.
Soit F la fonction où $F(x)$ désigne la fréquence des éléments de la série inférieurs ou égaux à x (F est la fonction de répartition c'est une fonction discontinue constante par intervalles, dans le cas d'une série discrète) la fonction quantile est la fonction Q définie sur [0 ; 1] par :
$Q(u) = \inf\{x, F(x) \geq u\}$. Ainsi Q (0,25) est le quartile 1, Q (0,75) est le quartile 3, Q (0,1) est le décile 1…

N.B.
– Dans ces définitions adoptées dans toutes les classes de lycée la médiane n'est pas toujours égale au quartile 2 : Q (0,5).
Une médiane est *un* nombre qui sépare la série en deux parties de même effectif. (On choisit de prendre la moyenne des deux valeurs centrales lorsque l'effectif de la série est impair).
Ce n'est pas toujours une valeur de la série contrairement aux quartiles, déciles, centiles.
Les valeurs de la médiane et des quartiles trouvées grâce aux définitions précédentes sont parfois différents de celles données par la calculatrice, elles-mêmes différentes de celles données par certains logiciels.
– Si les termes d'une série subissent une transformation affine (passage aux unités anglo-saxonnes, passage de francs en euros) les quantiles et la médiane subissent la même transformation.
– La moyenne et l'écart type sont sensibles aux valeurs extrêmes, la médiane et l'écart interquartile sont relativement plus stables. On utilise donc de plus en plus ce dernier couple pour résumer une série. Par exemple : si une seule valeur tend vers l'infini, la moyenne tend vers l'infini et la médiane reste constante.

Tige et feuille

En 1977, dans son livre *Exploratory Data Analysis*, John Tuckey a indiqué une méthode pour organiser ou représenter un ensemble de données numériques qu'il a appelées *stem and leaf* que l'on traduit en français par tige et feuille.
Il s'agit d'une disposition originale des données numériques qui peut être utile pour :
– dépouiller des données ;
– archiver ces données ;
– donner une représentation semi-graphique de ces données.

Expliquons la méthode sur un exemple :
Soit l'ensemble de données suivant qui représente l'ensemble des tailles, en centimètres, des élèves d'une classe terminale de lycée.

145	172	161	170	155	161	169	154	168	172
161	175	155	158	156	156	170	173	183	170
197	178	175	165	160	160	167	161	165	181

Nous allons ranger ces nombres en deux temps :
1. Chaque nombre est décomposé en deux parties :

145	→	14	et	5
		partie principale		feuille

On range sur une même ligne (appelée tige) tous les nombres ayant la même partie principale (*starting part*). La partie principale est inscrite une seule fois en début de ligne, les feuilles sont inscrites au fur et à mesure du dépouillement. On obtient ainsi :

14*	5
15*	5 4 5 8 6 6
16*	1 1 9 8 1 5 0 0 7 1 5
17*	2 0 2 5 0 3 0 8 5
18*	3 1
19*	7

On sépare les parties principales des feuilles par une ligne verticale.

2. On réécrit le tableau en rangeant, sur chaque tige, les feuilles dans l'ordre croissant, d'où la disposition finale :

14*	5
15*	4 5 5 6 6 8
16*	0 0 1 1 1 1 5 5 7 8 9
17*	0 0 0 2 2 3 5 5 8
18*	1 3
19*	7

On remarque que l'on obtient ainsi une sorte de diagramme en bâtons horizontal de la série statistique.

Avantages de la méthode tige et feuille

C'est une méthode pratique, simple et efficace pour réaliser un dépouillement en ordonnant les données.
Elle donne rapidement et sans outil de dessin un aperçu graphique d'une série statistique sans perte d'information sur les valeurs numériques.
Cette méthode permet d'obtenir facilement les quartiles d'une série statistique.

Médiane

La série précédente comporte 30 valeurs, une médiane est donc la demi-somme des valeurs des 15e et 16e individus, ces valeurs étant rangées dans l'ordre croissant. Il suffit de compter, à partir du haut du tige et feuille, les 15e et 16e feuilles. Ce sont 7 et 8 sur la tige 16, d'où la médiane **M_e = 167,5 cm**.

Quartiles d'ordre 1 et 3

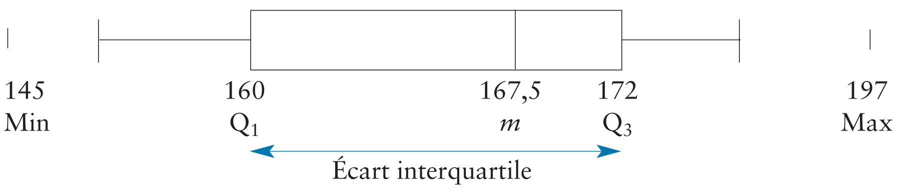
Il suffit de compter 8 feuilles à partir du haut du tige et feuille et 8 feuilles à partir du bas du tige et feuille : **Q_1 = 160 cm, Q_3 = 172 cm**.
Cette méthode peut être utilisée pour comparer visuellement deux séries statistiques : dans la précédente, les 12 premières valeurs correspondent aux tailles des filles de la classe, les autres valeurs correspondent, bien sûr, aux tailles des garçons

N.B. – On peut ajouter sur le diagramme certains renseignements statistiques :
– les effectifs cumulés croissants ;
– une marque indiquant la classe qui contient la médiane ;
– un résumé en cinq nombres de la série statistique.

1	14*	5
7	15*	4 5 5 6 6 8
(18)	16*	0 1 1 1 1 5 5 7 8
27	17*	9
29	18*	0 0 0 2 2 3 5 5 8
30	19*	8
		1 3
		7

160 — 167,5 — 172
145 — 197
Parenthèses indiquant la tige contenant la médiane
Q_1 — médiane — Q_3
minimum — maximum

Ce résumé en cinq nombres permet de faire une représentation graphique en diagramme en boîte (échelle non respectée).



– La méthode tige et feuille peut être également utilisée pour comparer visuellement deux séries statistiques.
Imaginons que dans la série étudiée les 12 premières valeurs correspondent aux tailles des filles de la classe, les autres valeurs correspondront donc, bien sûr, aux tailles des garçons.
On peut alors construire le double tige et feuille suivant :

5	14*	
4 5	15*	5 8 6 6
1 8 9 1 1	16*	5 0 0 7 1 5
5 2 0 2	17*	0 3 0 8 5
	18*	3 1
	19*	7

– Il existe plusieurs variantes du tige et feuille selon l'ordre de grandeur des données à traiter ainsi que la plus ou moins grande dispersion des données :
Par exemple, si l'on a la série suivante :

1 453	1 128	937	472	1 175	1 017	1 215	1 183
1 036	1 260	1 151	1 238	1 189	1 310	1 247	

On prendra des feuilles à deux chiffres :

4**	72
5**	
6**	
7**	
8**	
9**	37
10**	17 36
11**	28 51 75 83 89
12**	15 38 47 60
13**	10
14**	53

On constate que la valeur 472 se « détache » du reste de la série.
On peut aussi arrondir à la dizaine tous les nombres et représenter alors le nombre de dizaines (ce qui revient à changer d'unité). Il y a dans ce cas une légère perte d'information.

Unité 10

4*	7
5*	
6*	
7*	
8*	
9*	4
10*	2 4
11*	3 5 8 8 9
12*	2 4 5 6
13*	1
14*	5

L'ajustement affine

De quoi s'agit-il ?

Une situation statistique

Commençons par des exemples. Premier exemple : on étudie la morphologie des Français adultes et on s'intéresse à leur taille et à leur poids. Second exemple : en vue de détecter une maladie on regarde la concentration dans le sang d'une substance donnée, dans une population d'hommes sains, mais celle-ci varie avec l'âge des intéressés.
Pour caractériser les populations étudiées, il n'y a pas d'autre moyen que d'effectuer des mesures sur des individus membres de ces populations. Si tous les individus son interrogés, on a affaire à un recensement, sinon on prélève un échantillon. Intuitivement on voit bien que les personnes grandes ont tendance à avoir un poids plus élevé que les petites, que les hommes plus âgés ont une concentration plus élevée de la substance que les plus jeunes, mais il y a des variations. Pour une taille donnée les poids varient et la même situation existe pour le second exemple.
On est bien dans une situation relevant de la statistique. On veut caractériser une population à partir de mesures effectuées sur les individus qui la composent et il y a des variations dans la relation étudiée d'un individu à l'autre, mais ce ne sont pas les individus qui nous intéressent mais la population.

Recherche d'une relation

Sur chaque individu de la population on mesure donc deux variables ; appelons les $\boldsymbol{x}$ et $\boldsymbol{y}$, x_i et y_i seront les valeurs prises par ces variables pour l'individu i. On cherche s'il existe une relation simple, même approximative, entre ces deux variables. La plus simple serait la relation affine avec $\boldsymbol{y} = a\boldsymbol{x} + b$. Dans ce cas, les deux paramètres a et b suffisent à caractériser la population. On a donc construit un résumé des données qui permet d'oublier les $2n$ mesures faites sur les n individus.
Pour savoir s'il est possible de postuler qu'il existe entre les deux variables une relation affine approximative, il est pertinent de procéder à une représentation graphique. Dans le plan E^2, espace euclidien à deux dimensions, on considère le nuage des points de coordonnées (x_i, y_i), i variant de 1 à n. Chaque point représente un individu et le nuage la population ou au moins l'échantillon. Examiner à l'œil le nuage, c'est déjà procéder à un résumé ; on oublie les individus mais on réfléchit globalement à la population.
Si, en gros, le nuage ne s'écarte pas trop d'une droite, alors on peut envisager entre les deux variables une relation affine, reste à la trouver et à caractériser son intensité c'est-à-dire la qualité de la représentation.

Déterminer les paramètres de la relation

L'ajustement par les moindres carrés, *x* explique *y*

Jusqu'ici les deux variables jouent un rôle symétrique, mais dans beaucoup de problèmes, elles ne jouent pas le même rôle. L'une d'entre elles, par exemple $\boldsymbol{x}$, sert à prédire l'autre, elle est dite alors variable explicative, l'autre étant la variable à expliquer. Par exemple, l'âge sert à prévoir la composition sanguine et non l'inverse. On pose alors : $y = ax + b + \varepsilon$, où ε est un terme rendant compte des écarts existants entre la relation postulée et les constats expérimentaux. Pour l'individu i, on a : $y_i = ax_i + b + \varepsilon_i$. Dans ce contexte les coefficients a et b qui sont les plus adéquats sont ceux qui minimisent la somme des erreurs. Dans un cadre de description euclidien, il faudra minimiser la somme suivante : $\Sigma\varepsilon_i^2$.
La solution du problème est simple et les formules explicites permettant de calculer a et b optimaux à partir des observations bien connues ; le calcul est facilement automatisable et la plupart des calculatrices possèdent des touches spécialisées pour une exécution simple.

La droite représentative correspondante dans le plan E^2 se trace facilement, on montre qu'elle passe par le point G, centre de gravité du nuage de coordonnées (M(x), M(y)) où la fonction M désigne la moyenne arithmétique des valeurs prises par la variable considérée.
On peut aussi donner une autre interprétation géométrique de l'ajustement. Pour la précédente, dans E^2, les variables sont représentées par les axes et les individus par les points. On peut inverser les rôles et représenter, dans E^n, les variables par des points. La variable $\boldsymbol{x}$ sera représentée par le point X dont les coordonnées sont les valeurs prises par x, soit $x_1, x_2, \ldots, x_n$. Si $n = 3$, on peut même faire un dessin. Dans cette espace, on considère les vecteurs X – M(x) et Y – M(y) avec des notations évidentes : M(x) est le point dont toutes les coordonnées sont égales à M(x) et on identifie le point et le vecteur dont l'origine est le point de coordonnées (0, 0, …, 0) et l'extrémité le point X.
Expliquer la variable $\boldsymbol{y}$ par la variable $\boldsymbol{x}$ revient alors à rechercher le vecteur colinéaire à X – M(x) le plus proche du vecteur Y – M(y) et donc à projeter orthogonalement le dernier sur le premier, car s'il existe une relation affine entre x et y, alors les deux vecteurs précédents sont colinéaires. Le coefficient a est alors le rapport entre cette projection et le vecteur X – M(x). Si on a : $y = ax + b$, alors par linéarité, $M(y) = aM(x) + b$ et donc $Y - M(y) = a(X - M(x))$. La représentation affine est d'autant meilleure que l'angle entre les deux vecteurs est petit ou proche d'un angle plat. On montre facilement que le coefficient de corrélation linéaire noté ρ est le cosinus de cet angle, ρ proche de 1 correspond donc à un vecteur Y – M(y) proche de son projeté dans E^n et dans E^2 à un nuage de points proche de la droite représentative de l'ajustement qui a une pente positive. Si $\rho = -1$, on laisse au lecteur le soin de faire des commentaires analogues.

Et si *y* explique *x*

On pourrait croire que l'on a trouvé ainsi une relation affine entre deux variables. Si on considère une valeur non observée de $\boldsymbol{x}$, soit x_0, alors on peut prévoir la valeur correspondante de $\boldsymbol{y}$ soit : y_0 avec $y_0 = ax_0 + b$, cela correspond au modèle précédent. Il est alors tentant, avec la même relation, de vouloir prévoir la valeur prise par $\boldsymbol{x}$ pour une valeur donnée de $\boldsymbol{y}$. Malheureusement il n'en est rien car alors on ne respecte pas les présupposés du modèle précédent, on explique alors x par y donc on écrit, pour l'individu i : $x_i = cy_i + d + \eta_i$
et on cherche c et d qui minimisent la quantité : $\Sigma\eta^2_i$
On a une solution telle que c est différent de $\frac{1}{a}$ et d différent de $-\frac{b}{a}$.

Dans la représentation dans E^2, la droite représentant le nouvel ajustement est différente de la première en général ; elles passent toutes les deux par le centre de gravité G du nuage. Dans la représentation dans E^n on projette orthogonalement le vecteur X – M(x) sur le vecteur Y – M(y) et non l'inverse.

N.B. – À noter que l'erreur est fréquente, elle peut exister dans certains sujets d'examen passés voire dans des manuels.

D'autres ajustements

Avant l'arrivée des calculatrices, la recherche des droites des moindres carrés nécessitait des calculs à la main longs et pénibles, alors on s'affranchissait du modèle précédent et on cherchait grossièrement à remplacer dans E^2 le nuage de points par une droite. La méthode la plus connue est celle de la droite de Mayer. On divise le nuage en deux sous-nuages. Le premier comprend les points dont l'abscisse est inférieure à la médiane des x, le second les autres. Soient G_1 et G_2 les centres de gravité de chacun des sous-nuages et on prend comme droite ajustant « au mieux » le nuage la droite (G_1G_2), elle passe par le point G. Dans certains cas la méthode ne donne pas de mauvais résultats, en particulier quand il y a une forte corrélation entre les deux variables, mais il peut y avoir des déconvenues. Surtout elle ne repose sur aucun modèle. La méthode a été très populaire parmi les professeurs de mathématiques car elle permettait d'illustrer les propriétés des barycentres, mais cela n'a rien à voir avec la statistique.

Quant à utiliser des méthodes empiriques, on peut aussi tracer dans le nuage de points la droite qui, à l'œil, s'ajuste au mieux. L'expérience montre que l'on est alors proche d'une autre droite explicitée ci-dessous.
Si introduire une dissymétrie entre les variables n'est pas pertinent, on peut chercher la droite la plus proche du nuage au sens suivant : on minimise la somme des carrés des distances des points du nuage à la dite droite et on remplace le nuage à deux dimensions par un nuage à une dimension. On projette alors orthogonalement les points du nuage sur cette droite et on obtient le nuage à une dimension le plus proche du nuage à deux dimensions. Cette technique n'est utilisée concrètement que quand on a plus de deux variables et que l'on cherche un nuage à une, deux dimensions, le plus proche d'un nuage à k dimensions.

Interpréter un ajustement

Trouver une relation affine approximative entre deux variables, l'une expliquant l'autre, ne veut pas dire qu'il y a une relation de causalité entre les deux. Les exemples sont nombreux, surtout quand le temps intervient. Ainsi on a observé en Allemagne une bonne corrélation entre le nombre des cigognes arrivant une année donnée et le nombre des naissances de la même année. La population considérée est celle des années de la période étudiée, x est le nombre de cigognes arrivées et y le nombre de naissances. L'ajustement est très bon, d'ailleurs il est bien connu que les cigognes apportent les bébés !

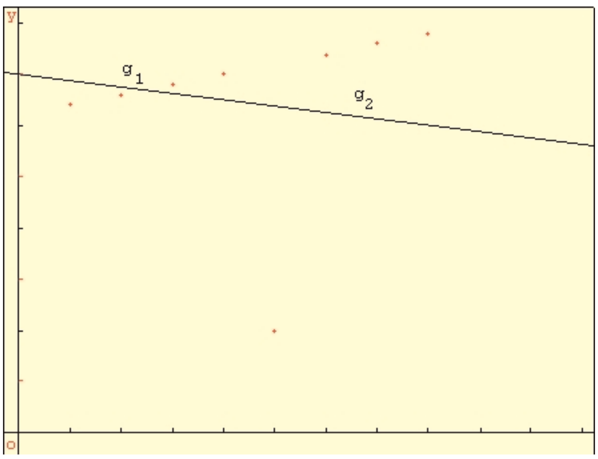
Exemple de recherche d'une droite passant au plus « près » d'un nuage de points

Une personne a relevé sa température à intervalles de temps réguliers elle a trouvé les résultats suivants :

Temps	1	2	3	4	5	6	7	8
Température	37,2	37,3	37,4	37,5	35	37,7	37,8	37,9

Une représentation de nuage de points permet de penser que la température augmente régulièrement et donc qu'une bonne approximation affine sera obtenue par une droite à coefficient directeur positif.

Mais la droite obtenue par la méthode de Mayer nous donne un coefficient négatif !
Les points moyens des deux sous-nuages sont : g_1 (2.5 ; 37,35) et g_2 (6,5 ; 37,1). Le cœfficient directeur est : $-\frac{0,25}{4} = -0,0625$. (La droite obtenue par la méthode de Mayer est sensible aux valeurs extrêmes : normal puisque la moyenne l'est aussi.)



Loi de probabilité sur un ensemble fini

On introduira le concept de probabilité en vue de modéliser des situations particulières pour lesquelles aucune propriété de symétrie ne permet de parler « d'égalité des chances ».

– Exemple : on dispose du tableau suivant donnant les occurrences de différents groupes sanguins sur une série de 10 000 naissances dans des maternités de France.

A	*B*	*AB*	*O*
4 546	863	449	4 142

Modéliser cette situation, c'est définir une loi de probabilité sur E = {*A*, *B*, *AB*, *O*}, c'est-à-dire affecter à chaque élément de E un nombre (sa probabilité) positif ou nul, la somme de ces nombres valant 1.
On admettra la loi des grands nombres ; c'est un théorème de mathématiques qui peut être énoncé sous une forme vulgarisée : « Si on choisit n éléments selon une loi de probabilité P, indépendamment les uns des autres, alors la distribution des fréquences est voisine de P lorsque n est grand. »
Les élèves ont observé en seconde la diminution de l'ampleur de la fluctuation d'échantillonnage avec la taille de la série de données : il s'agissait de l'observation d'une loi empirique. La correspondance entre loi empirique et théorème mathématique est un des fondements de la modélisation par des modèles probabilistes.
En vertu de la loi des grands nombres, on choisira le modèle défini par
$P = \frac{4546}{10000}, \frac{863}{10000}, \frac{449}{10000}, \frac{4112}{10000}$ ou la loi P' = (0,45 ; 0,09 ; 0,05 ; 0,41). Dans le modèle défini par P', on dira qu'un bébé qui va naître a une probabilité 0,41 d'être du groupe O. On indiquera qu'une validation d'un modèle consisterait par exemple à prendre de nouvelles données expérimentales et à comparer ce qu'on obtient et ce que prévoit le modèle (compte tenu de la fluctuation d'échantillonnage, les modalités de cette comparaison ne sont pas simples et sont hors programme).
La liste des groupes sanguins de 10 bébés (en excluant les naissances multiples !) sera alors matérialisée comme la liste des résultats de 10 expériences « identiques et indépendantes » : la probabilité de la liste obtenue est égale au produit des probabilités de chacun de ses termes.
On veut maintenant faire un modèle tenant compte du facteur rhésus ; le tableau des données est le suivant :

	A	*B*	*AB*	*O*
RH+	968	753	385	3 566
RH-	578	110	64	576

La loi considérée ici sera une loi de probabilité sur {A+, A-, B+, B-, AB+, AB-, O+, O-}. On pourra alors calculer la probabilité de chaque groupe, celle d'être rhésus positif, la probabilité de l'être sachant O, etc.

Tests de dépistage

Pour diagnostiquer une maladie ou dépister un caractère, chez des personnes ou des animaux, on utilise des tests cliniques ou biologiques.
Il y a deux phases dans l'élaboration d'un test :
– l'étalonnage à l'aide d'un groupe témoin constitué de personnes ou d'animaux (que nous qualifierons de sujets) dont on connaît le caractère (par exemple, avoir ou non la maladie) ;
– l'utilisation où l'on applique ce test pour détecter la maladie.

Pour simplifier, par la suite, on notera T : « le test est positif » et M : « le sujet est malade ». On cherche à établir un lien entre ces deux événements.

Dans la phase d'étalonnage du test, on s'intéresse à deux indices : la probabilité d'avoir un test positif sachant que le sujet est malade et la probabilité que le test soit négatif sachant que le sujet n'est pas malade.
Selon certains auteurs, on parle de :
– la sensibilité du test : $S_e = p_M(T)$, encore notée $p(T/M)$;

– la spécificité du test : $S_p = p_{\overline{M}}(\overline{T})$, encore notée $p(\overline{T}/\overline{M})$.
Ces indices sont indépendants de la probabilité *a priori* d'être malade, souvent appelée prévalence de la maladie dans la population étudiée. Pour que le test soit fiable, les valeurs de $p_M(T)$ et $p_{\overline{M}}(\overline{T})$ doivent être les plus proches possible de 1. Cependant il ne faut pas perdre de vue que ces deux probabilités sont déterminées à partir de statistiques, ce qui peut poser problème.

Dans la phase d'utilisation du test, on cherche à détecter les personnes malades. La difficulté est que le test peut être positif à tort (le test est positif alors que la personne n'est pas malade) ou négatif à tort (le test est négatif alors que la personne est malade). Le clinicien qui désire poser un diagnostic peut s'intéresser, entre autres, à deux paramètres :
– la valeur prédictive positive : $VPP = p_T(M)$;
– la valeur prédictive négative : $VPN = p_{\overline{T}}(\overline{M})$.

On a accès aux probabilités $p_T(\overline{M})$ et « inverses » des probabilités précédentes, $p_M(T)$ et $p_{\overline{M}}(M)$, en utilisant la formule des probabilités totales :

$$p(T) = p_M(T) \times p(M) + p_{\overline{M}}(T) \times p(M)$$

d'où l'on tire :

$$p(T) = p_M(T) \times p(M) + (1 - p_{\overline{M}}(\overline{T})) \times (1 - p(M)$$

$$p(T) = S_e \times p(M) + (1 - S_p) \times (1 - p(M)$$

et en utilisant les relations :

$$p_T(M) = p_M(T) \times \frac{p(M)}{p(T)}$$

$$p_T(\overline{M}) = p_M(\overline{T}) \times \frac{1 - p(M)}{1 - p(T)}$$

on retrouve l'expression donnée par le théorème de Bayes :

$$p_T(M) = \frac{p_M(T) \times p(M)}{p_M(T) \times p(M) + p_{\overline{M}}(T) \times p(\overline{M})}$$

Cependant, concrètement, les probabilités qui interviennent dans cette formule sont estimées à partir de statistiques. Il y a lieu de s'assurer que le modèle qui en résulte est adapté à la population étudiée.
En remplaçant $p_M(T)$ par S_e, $p(\overline{M})$, par $1 - p(M)$, $p_{\overline{M}}(T)$ par $1 - p_{\overline{M}}(\overline{T})$ et $p_{\overline{M}}(\overline{T})$ par S_p on obtient :

$$VPP = \frac{S_e \times p(M)}{S_e \times p(M) + (1 - S_p) \times (1 - p(M))}$$

On obtiendrait de même :

$$VPN = \frac{S_p \times (1 - p(M)}{(1 - S_e) \times p(M) + S_p \times (1 - p(M))}$$

On remarque que, contrairement à S_e et S_p, les valeurs prédictives VPP et VPN dépendent de la probabilité *a priori* d'être malade. On est donc amené à étudier VPP et VPN en fonction de la variable $p(M)$.

Par ailleurs, l'efficacité du test peut être mesurée par $p(M \cap T) + p(\overline{M} \cap \overline{T})$, que certains auteurs notent E_f : $E_f = p(M \cap T) + p(\overline{M} \cap \overline{T})$.
On peut encore écrire : $E_f = S_e \times p(M) + S_p \times (1 - p(M))$.

Exemples « mathématiques – biologie »

La génétique vue par les biologistes

Définition et localisation du programme génétique

Le programme génétique est situé dans chaque cellule de l'organisme. Il correspond au plan de fabrication et fonctionnement de cette cellule et par conséquent de l'organisme.

Ce programme se situe dans le noyau, au niveau :
– de la chromatine lorsque la cellule ne se divise pas ;
– des chromosomes quand la cellule se divise.
Chromosomes et chromatine correspondent au même matériel génétique (ADN) mais à des degrés différents de concentration.
L'ADN est la molécule porteuse du programme génétique. Chaque caractère héréditaire est codé par une portion précise de cette molécule : le gène. Un gène existe sous différentes versions appelées : allèles. Ceux-ci sont responsables des différentes variantes d'un caractère héréditaire.

Dans l'exercice proposé, on croise des drosophiles femelles de lignée pure (homozygotes) aux yeux rouges et aux ailes entières, avec des drosophiles mâles de lignée pure aux yeux marron et aux ailes échancrées.
La première génération (F1) ne donne que des drosophiles aux yeux rouges et aux ailes entières.
Lorsque l'on croise des drosophiles femelles de la première génération (F1) avec des mâles de lignée pure aux yeux marron et aux ailes échancrées, on obtient :
– 410 drosophiles aux yeux rouges et aux ailes entières ;
– 400 drosophiles aux yeux marron et aux ailes échancrées ;
– 111 drosophiles aux yeux rouges et aux ailes échancrées ;
– 109 drosophiles aux yeux marron et aux ailes entières.

On note les événements suivants :
– R : « la drosophile a les yeux rouges » ;
– E : « la drosophile a les ailes entières » ;
– m : « la drosophile a les yeux marron » ;
– e : « la drosophile a les ailes échancrées ».

N.B. – Il n'y a jamais de *crossing-over* chez le mâle de la drosophile.

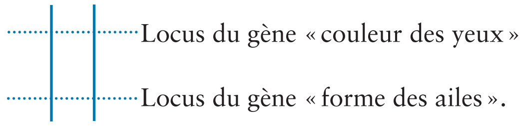
1. Que peut-on déduire du fait que le premier croisement ne donne que des drosophiles aux yeux rouges et aux ailes entières ?
2. Calculer les probabilités d'apparition de chacun de ces phénotypes.
3. Construire l'échiquier de croisement correspondant au deuxième croisement.
4. En déduire les probabilités des différents phénotypes du croisement F2 de drosophiles issus de F1.
5. On examine l'une des drosophiles de F2, prise au hasard. Quelle est la probabilité pour qu'elle ait les yeux rouges ? Quelle est la probabilité qu'elle ait les yeux marron sachant qu'elle a les ailes échancrées ? Quelle est la probabilité d'avoir un individu hybride ?

Caractères héréditaires → Variantes
1) couleur des yeux → rouge, marron
2) forme des ailes → entières, échancrées

Déterminisme génétique : Gène — Allèles

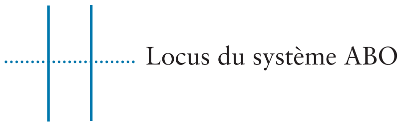
Pour les organismes diploïdes, les chromosomes sont en nombre pair dans leurs cellules (par exemple $2n = 46$ pour l'espèce humaine ; $2n = 8$ pour la drosophile). Ces chromosomes sont identiques deux à deux en forme, taille… Ce sont des chromosomes homologues. Ainsi pour l'espèce humaine, il y a 23 paires de chromosomes homologues ; pour la drosophile, 4 paires de chromosomes homologues.
Les gènes ont une localisation précise sur les chromosomes : le locus du gène, partagé par les deux chromosomes homologues.
Certains gènes sont portés sur la même paire de chromosomes, ce sont des gènes liés.

Dans l’exercice proposé :

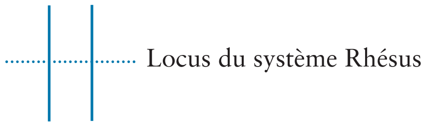


1 paire de chromosomes homologues

D’autres gènes sont portés par des paires différentes de chromosomes homologues, ce sont des gènes libres.
Exemple groupes sanguins : système ABO et système Rhésus.



Paire n° 1 de chromosomes homologues



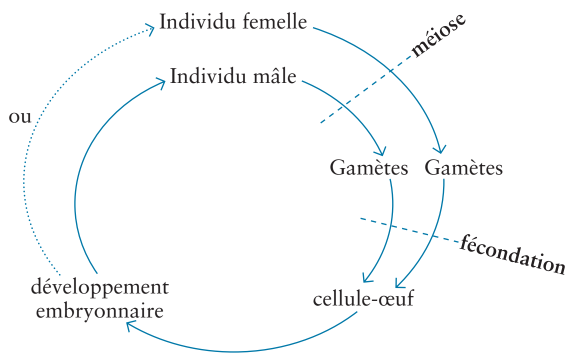
Paire n° 2 de chromosomes homologues

Les allèles portés par les chromosomes homologues peuvent être identiques (homozygotes) ou différents (hétérozygotes). On appelle génotype l’ensemble des allèles présents dans la cellule. On le représente par une « fraction » symbolisant une paire de chromosomes homologues.
Exemples de génotypes homozygotes :
– pour la couleur des yeux : marron /marron ; rouge / rouge ;
– pour les groupes sanguins : A/O ; A/B…
Exemples de génotypes hétérozygotes :
– pour la couleur des yeux : marron/rouge ;
– pour les groupes sanguins : A/O ; A/B…

Le génotype détermine le phénotype, c’est-à-dire les structures, fonctionnements… observables. Le phénotype est l’expression du génotype, mais il faut alors tenir compte des relations de dominance entre allèles pour les génotypes hétérozygotes :
Par exemple, pour les drosophiles, le génotype (marron / rouge) s’exprime par un phénotype « yeux rouges », donc seul l’allèle « rouge » est exprimé. Il est dit dominant et l’allèle « marron » est alors récessif.
Pour les groupes sanguins, le génotype A/O correspond à un phénotype « A » : l’allèle A est dominant et l’allèle O récessif.
Il peut y avoir expression des deux allèles : ils sont alors codominants : c’est le cas des allèles A et B des groupes sanguins ; le phénotype « AB » correspond au génotype A/B.

Transmission des allèles au cours de la reproduction sexuée :
Le cycle biologique des espèces est caractérisé par deux phénomènes complémentaires :
– la méiose, intervenant au cours de la formation des gamètes (cellules reproductrices) ;
– la fécondation, fusion de deux gamètes, formant la cellule-œuf, origine d’un nouvel individu.

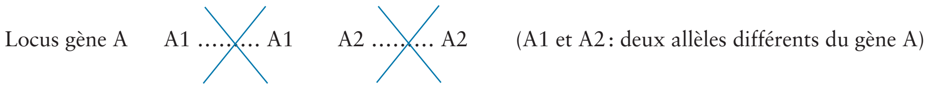
Cycle biologique d’une espèce diploïde :



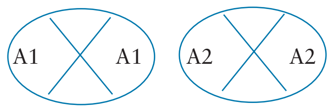
Pour une espèce diploïde, la cellule-œuf et les cellules des individus adultes, sont à $2n$ chromosomes. Les gamètes ne possèdent donc que n chromosomes (cellules *haploïdes*) : ils n'ont qu'un représentant de chaque catégorie de chromosome.
Donc la méiose s'accompagne d'une réduction de moitié du nombre de chromosomes, les chromosomes homologues se séparent. La fécondation correspond à un retour à la diploïdie : chaque gamète apportant un exemplaire de chaque type chromosomique, la cellule-œuf comporte donc les chromosomes homologues.
Les chromosomes homologues portant des allèles différents pour une majorité des gènes, ces deux phénomènes s'accompagnent donc d'un brassage allénique. Pour qu'il y ait brassage allénique les génotypes parentaux doivent être hétérozygotes.

Les brassages des allèles au cours de la méiose

La méiose est constituée de deux divisions cellulaires successives. Au départ, les $2n$ chromosomes sont dupliqués (deux chromatides identiques génétiquement, mais les allèles peuvent être différents sur chaque chromosome homologue).



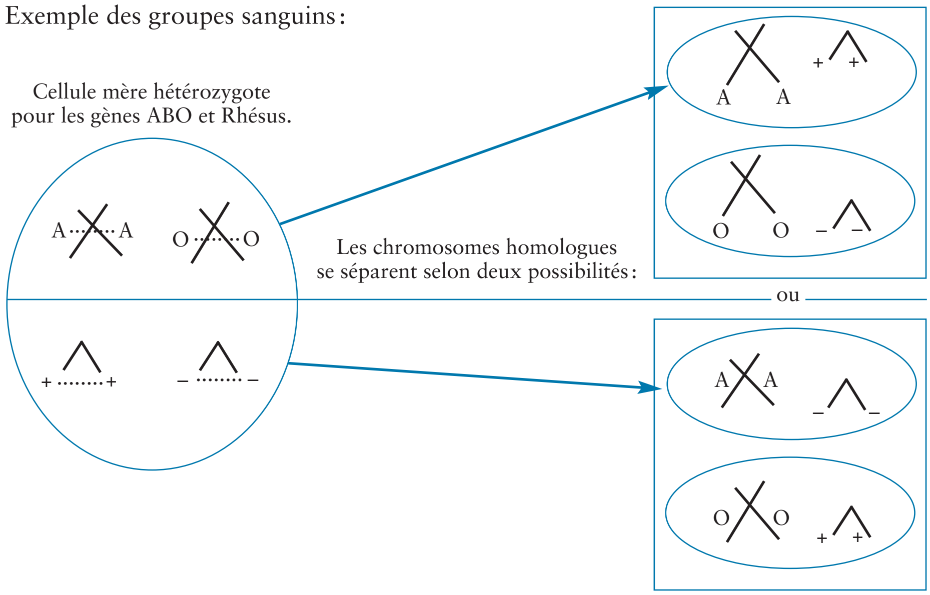
Ces deux chromosomes homologues se séparent au cours de la 1re division de méiose. On obtient deux cellules filles à n chromosomes 2 chromatides :



On observe aussi que les deux cellules possèdent un programme génétique différent, il y a eu séparation des allèles au cours de la séparation des chromosomes homologues.

Ce phénomène est amplifié si on considère la transmission de plusieurs gènes :

Exemple des groupes sanguins :



Cellules filles obtenues
à l'issue de la 1re division
de méiose.

Ce brassage, résultat de la ségrégation indépendante des allèles au cours de la séparation des chromosomes homologues est dit brassage inter-chromosomique.
Il forme, pour n gènes libres, 2^n gamètes génétiquement différents et équiprobables.

Au cours de cette 1re division des échanges de fragments de chromatides peuvent se produire : ce sont les *crossing-over*. On obtient alors des chromatides de programmes génétiques différents.

Exemple : drosophiles de l'exercice : hétérozygotes pour les deux gènes (R = yeux rouges ; m = yeux marron ; E = ailes entières ; é = ailes échancrées).

R m R m locus gène « couleur des yeux » R ...... R m ... m
E é E é locus gène « forme des ailes » E...... é E...... é
Entrecroisement des chromosomes homologues Résultat : le *crossing-over*.

Ce brassage intervenant entre chromosomes homologues d'une même paire est appelé brassage intra-chromosomique. Après séparation des chromosomes homologues, on obtient quatre gamètes génétiquement différents : (R, E) ; (R, é) ; (m, E) ; (m, é).
Le *crossing-over* intervient rarement au cours de la méiose, les gamètes issus de ce brassage sont donc en quantité inférieure par rapport à ceux issus des méioses sans *crossing-over* : (R, E) et (m, é). On a donc des gamètes dits parentaux (R, E) ; (m, é) qui sont plus abondants que les gamètes dits recombinés (R, é) ; (m, E). Le brassage intra-chromosomique est ensuite amplifié par le brassage interchromosomique.

La 2[e] division de méiose se caractérise par une séparation des chromatides. On obtient donc des cellules à *n* chromosomes à une chromatide.

Le brassage des allèles au cours de la fécondation

Lors de la fusion des gamètes, les chromosomes homologues se retrouvent regroupés dans une même cellule. Les brassages de la méiose ont créé des génotypes haploïdes originaux dans les gamètes, leur regroupement amplifie cette création de génotypes diploïdes nouveaux.
On réunit tous ces génotypes dans un tableau à double entrée appelé « tableau de rencontre des gamètes » ou « échiquier de croisement ».

– Exemple 1 : groupes sanguins ABO et rhésus.
La mère est hétérozygote A/O ; Rh + /Rh-. Par brassage inter-chromosomique, elle peut former 4 catégories de gamètes génétiquement différents, équiprobables.
Le père est hétérozygote B/O ; Rh+ / Rh-. Par brassage inter-chromosomique, il peut former 4 catégories de gamètes génétiquement différents, équiprobables.
On regroupe les gamètes dans le tableau, et on en déduit les génotypes possibles des descendants :

Gamètes de la mère / Gamètes du père	*A, +*	*A, -*	*O, +*	*O, -*
B, +	A/B ; + / +	A/B ; + / -	B/O ; + / +	B/O ; + / -
B, -	A/B, + / -	A/B ; -/ -	B/O ; + / -	B/O ; - / -
O, +	A/O, + / +	A/O ; + / -	O/O ; + / +	O/O ; + / -
O, -	A/O, + / -	A/O ; - / -	O/O ; + / -	O/O ; - / -

Chaque génotype est équiprobable avec une probabilité de : $\frac{1}{16}$.

– Exemple 2 : croisement proposé dans l'exercice « Exercice 1 commun mathématiques et SVT ».
Les gènes sont liés, il y a *crossing-over* chez la femelle.

La femelle F1 est hétérozygote (E,R / é,m). Par brassage intra-chromosomique et ségrégation des allèles, elle fournit quatre catégories de gamètes génétiquement différents, non équiprobables :
– (E, R) et (é, m) plus abondants ;
– (é, R) et (E,m) recombinés, issus du *crossing-over*, moins abondants.
Le mâle est homozygote é,m / é,m. Il ne fournit qu'une catégorie de gamètes : (é,m).

D'où l'échiquier de croisement :

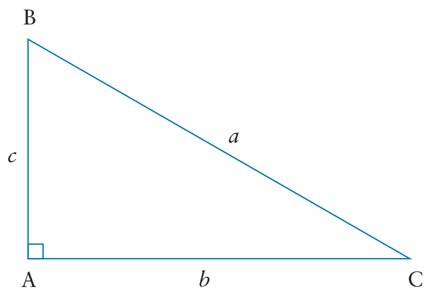
Gamètes de la femelle / Gamètes du mâle	E, R	é, m	E, m	é, R
é, m	E/é ; R/m	é/é ; m/m	E/é ; m/m	é/é ; R/m

Les génotypes obtenus ne sont pas équiprobables : les deux derniers sont moins représentés puisqu'ils sont issus de gamètes recombinés. Les génotypes s'expriment au niveau des phénotypes : donc les drosophiles à ailes entières et yeux rouges et celles à ailes échancrées et yeux marron sont les plus représentées dans la population obtenue à l'issue de ce croisement.

Les outils pour les sciences physiques et chimiques

Trigonométrie

Dans un triangle ABC rectangle en A, nous avons : $AB^2 + AC^2 = BC^2$.

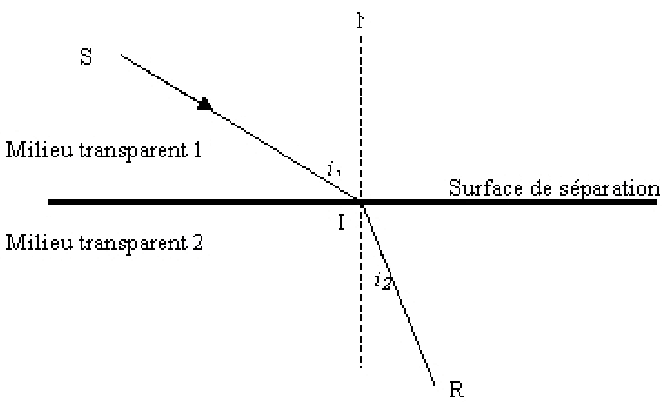


$\sin C = \dfrac{AB}{BC}$; $\cos C = \dfrac{AC}{BC}$; $\tan C = \dfrac{AB}{AC}$.

$\sin B = \dfrac{AC}{BC}$; $\cos B = \dfrac{AB}{BC}$; $\tan B = \dfrac{AC}{AB}$.

Loi de la réfraction

On appelle réfraction le changement de direction que subit la lumière à la traversée de la surface de séparation entre deux milieux transparents.
Pour étudier la réfraction, on fait arriver dans un milieu transparent 1 (l'air par exemple) un rayon lumineux sur la surface d'un milieu transparent 2 (l'eau par exemple) comme le montre le schéma ci-dessous.

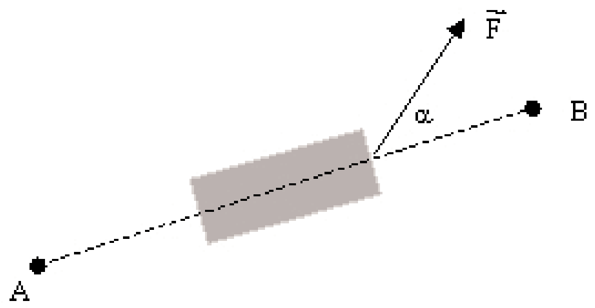


Le rayon incident (SI) donne naissance à un rayon réfracté (IR). La perpendiculaire (IN) au plan de séparation des deux milieux détermine un angle d'incidence i_1 et un angle de réfraction i_2.
La loi de Descartes concernant la réfraction est $n_1 \sin i_1 = n_2 \sin i_2$ où n_1 et n_2 représentent les indices de réfraction des milieux 1 et 2. L'indice de réfraction est une grandeur caractéristique d'un milieu transparent. Il s'exprime sans unité.

Travail et puissance d'une force constante de translation

Définition

Une force constante est une force dont la direction, le sens et l'intensité ne se modifient pas au cours du déplacement. La force est dite de translation si son point d'application parcourt une droite.
Le travail d'une force $\vec{F}$ constante dont le point d'application effectue un déplacement rectiligne AB = l, faisant un angle de mesure $\propto$ avec la direction de la force, est donné par la relation :
$W_{A \to B} = F \times l \times \cos \propto$, soit $\vec{F}.\vec{AB}$.



W s'exprime en joules (J) si F est en newtons (N), l en mètres (m) et α en degrés (°).
N. B. – Si la droite d'action de la force $\vec{F}$ est confondue avec la direction du déplacement, on a $\alpha = 0°$ ou $\propto = 180°$.
L'expression du travail s'écrit alors $W = F \times l$, car $\cos 0° = 1$ et $W = -F \times l$ car, $\cos 180° = -1$.

Expression du travail en fonction de la valeur de α

– Si $0° \leq \alpha < 90°$ alors la force $\vec{F}$ provoque le déplacement, elle est dite motrice. Le travail est positif, c'est un travail moteur.
– Si $90° < \alpha \leq 180°$ alors la force $\vec{F}$ s'oppose au déplacement, elle est dite résistante. Le travail est négatif, c'est un travail résistant.
– Pour $\alpha = 0°$, le travail est maximum : $W = F \times l$.
– Pour $\alpha = 90°$, le travail est nul : $W = 0$. Le travail d'une force perpendiculaire au déplacement de son point d'application est nul.
– Pour $\alpha = 180°$, le travail est $W = -F \times l$.

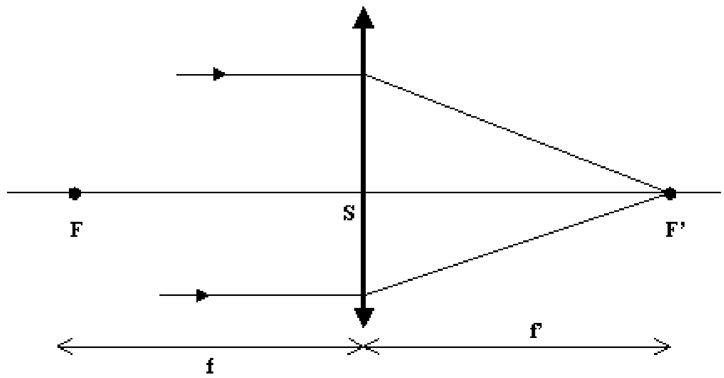
Optique géométrique

Calcul des combinaisons optiques : optique paraxiale – Conditions de Gauss[1]

Les méthodes de calcul des combinaisons optiques que nous allons étudier sont en principe limitées aux lentilles minces et aux faisceaux lumineux peu écartés de l'axe et dont l'angle d'ouverture est inférieur à 3°. Les rayons lumineux sont paraxiaux.
Ce sont les conditions d'approximation de Gauss. En pratique, la précision des calculs sera d'autant plus altérée que l'on s'écartera de ces limites.

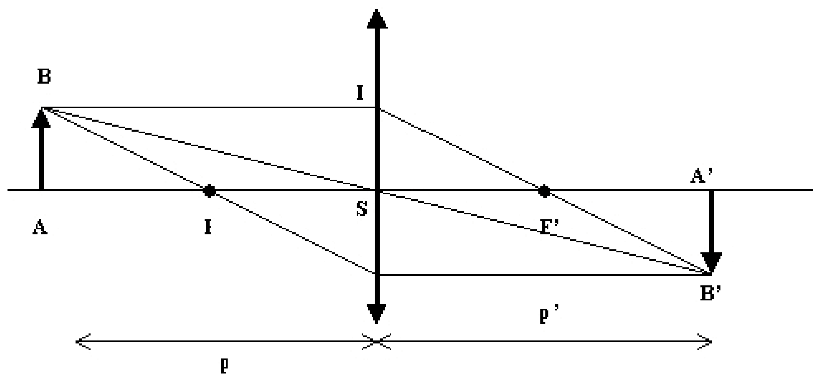
1. D'après le site http://serge.bertorello.free.fr

Conventions – La figure ci-dessous illustre les conventions que nous allons utiliser. Elle représente une lentille convergente. Celle-ci comporte deux foyers.



F est le foyer côté objet, nous l'appellerons foyer objet.
f est la distance focale objet.
F' est le foyer image.
f' est la distance focale image.
S est le centre de la lentille.

Position d'une image : détermination graphique



Considérons un objet matérialisé par la flèche « $\overrightarrow{AB}$ » et déterminons la position et la taille de son image. Pour cela, nous pouvons utiliser deux des trois rayons particuliers suivants :

– Le rayon incident [BI] parallèle à l'axe, qui émerge suivant [IF']. En effet, nous avons vu qu'un rayon incident parallèle à l'axe sort de la lentille en passant par le foyer image F'.
– Le rayon incident [BF], qui émerge parallèlement à l'axe puisqu'il passe par le foyer objet F. C'est le cas symétrique du précédent.
– Le rayon incident [BS], qui n'est pas dévié car il passe par le centre S de la lentille.

Cette méthode graphique nous permettra donc de concevoir très simplement des combinaisons optiques destinées à former des images.

Première formule de Descartes

Nous pouvons aussi déterminer la position de l'image par le calcul. Pour cela, René Descartes a démontré la formule suivante : $\frac{1}{p'} = \frac{1}{p} + \frac{1}{f'}$ avec p la distance entre l'objet et la lentille ; p' la distance entre la lentille et l'image et f' la distance focale de la lentille.

Taille de l'image

Méthode graphique

La construction graphique de la figure ci-dessus nous donne la taille de l'image, à condition de connaître l'échelle du schéma.

Deuxième formule de Descartes

En analysant la figure précédente, nous constatons que les triangles ABS et A′B′S sont semblables (programme de seconde). Par conséquent, nous pouvons énoncer la relation suivante : $g = \frac{A'B'}{AB} = \frac{A'S}{AS} = \frac{p'}{p}$ où g désigne le grandissement. La relation définissant g est $= \frac{f'}{f'+p}$.

Fonction logarithme décimal

La touche $\boxed{\text{log}}$ de la calculatrice permet d'obtenir le logarithme décimal de certains nombres.

Ainsi $\boxed{\text{log}}$ 2 $\boxed{\text{EXE}}$ donne 0,301 029

$\boxed{\text{2}^{\text{nd}}}$ $\boxed{\text{log}}$ 2 $\boxed{)}$ $\boxed{\text{ENTER =}}$ donne 10^2.

Compléter le tableau suivant :

a	0,001	0,01	0,1	1	10	100	1 000
log a							

Comparer les résultats précédents avec l'écriture des nombres décimaux précédents sous la forme 10^n pour n entier positif ou négatif.

Quelques autres propriétés : pour a et b réels strictement positifs :
$\log (a \times b) = \log a + \log b$. La fonction logarithme décimal transforme un produit en une somme.

$\log a^n = n\log a$ pour n réel ; $\log\left(\frac{a}{b}\right) = \log a - \log b$.

Recherche du pH

Le pH est la mesure de l'acidité d'une solution. Il signifie « potentiel d'hydrogène ». En 1909, S.P.L. Soerensen a défini l'acidité d'une solution comme étant l'opposé du logarithme décimal de la concentration en ions hydrogène. Dans cette notation, *p* est l'abréviation du mot allemand *potenz* (potentiel) et H est le symbole de l'hydrogène.

Le pH des solutions acides est inférieur à 7 et celui des solutions alcalines (ou basiques) est supérieur à 7. Quand une substance a un pH de 7, elle est neutre.
Comme l'échelle de pH est logarithmique, une différence d'une unité entraîne en réalité un décuplement (multiplication par 10). Par exemple, l'acidité d'un échantillon au pH de 5 est 10 fois supérieure à celle d'un échantillon au pH de 6. Un écart de deux unités, de 6 à 4, indique une acidité 100 fois supérieure.

Le pH au quotidien

Substance	PH
Acide chlorhydrique molaire	0
Drainage minier acide	< 1,0
Batterie acide	< 1,0
Acide gastrique	2,0
Jus de citron	2,4
Cola	2,5
Vinaigre	2,9
Jus d'orange ou de pomme	3,5
Bière	4,5
Café	5,0
Thé	5,5

Le pH au quotidien	
Substance	**PH**
Pluie acide	< 5,6
Lait	6,5
Eau pure	7,0
Salive humaine de	6,5 à 7,4
Sang de	7,34 à 7,45
Eau de mer	8,0
Savon de	9,0 à 10,0

Le pH est tel que $[H_3O^+] = 10^{-pH}$ ou $pH = -\log [H_3O^+] = -\log C$, où C désigne la concentration de la solution en ions H_3O^+.

Les constantes d'acidité K_a et p K_a avec p $K_a = -\log K_a$.

Lexique

Antécédent d'un nombre réel par une fonction numérique

Soit I un intervalle de **R**, f une fonction de I dans **R** et m un nombre réel. On appelle antécédent de m par f tout réel x appartenant à I tel que $f(x) = m$ (voir exercices 137 sur le site Euler).

Arrondi (ou valeur arrondie)

– Soit x un nombre, l'arrondi de x à l'unité est le nombre entier a tel que $a - 0{,}5 \leq x < a + 0{,}5$. Il s'agit d'une valeur approchée de x à 1 près.
– Soit x un nombre, l'arrondi de x au dixième est le nombre décimal a tel que $10a$ est entier et $a - 0{,}05 \leq x < a + 0{,}05$. Il s'agit d'une valeur approchée de x à 0,1 près.
– Soit x un nombre, l'arrondi de x au centième est le nombre décimal a tel que $100a$ est entier et $a - 0{,}005 \leq x < a + 0{,}005$. Il s'agit d'une valeur approchée de x à 0,01 près.
– Soit x un nombre, l'arrondi de x au millième est le nombre décimal a tel que $1000a$ est entier et $a - 0{,}0005 \leq x < a + 0{,}0005$. Il s'agit d'une valeur approchée de x à 0,001 près.
– Soit x un nombre réel et n un entier naturel. On appelle arrondi (ou valeur arrondie) de x à 10^{-n} le nombre décimal a tel que $a \times 10^{n}$ est entier et $a - 5 \times 10^{-n-1} \leq x < a + 5 \times 10^{-n-1}$

Classe modale

Soit S une série statistique quantitative continue à une variable dont les modalités sont regroupées par classes de même amplitude.
Une classe modale est une classe de plus grand effectif de cette série.

Coefficient directeur d'une droite

Soit D une droite du plan non parallèle à l'axe des ordonnées. Il existe un unique couple de réels (m, p) tel que la droite D ait pour équation réduite $y = mx + p$.
Le réel m est appelé coefficient directeur de la droite D.

Coefficient multiplicateur ou multiplicatif

Si p désigne un nombre réel positif ou nul et x une variable réelle positive. x subit un pourcentage d'évolution égal à p % à partir d'une valeur x_0, sa nouvelle valeur x_1 est $x_0 + \frac{p}{100}x_0$ ou $\left(1 + \frac{p}{100}\right)x_0$.

Le nombre $k = 1 + \frac{p}{100}$ est appelé **coefficient multiplicateur** (**ou multiplicatif**) de la variable x.
Si x subit des évolutions successives de coefficients multiplicateurs k_1 et k_2 alors x subira une évolution de coefficient multiplicateur $k_1\,k_2$.

Courbe représentative d'une fonction numérique

Soit I un intervalle de **R**, f une fonction de I dans **R**. On appelle courbe représentative de la fonction f dans le plan muni d'un repère $(O\,;\ \vec{i},\ \vec{j})$ l'ensemble des points M du plan pour lesquels il existe x appartenant à I tel que M ait pour coordonnées $(x, f(x))$ où x décrit l'intervalle I.

Droite d'ajustement affine par la méthode des moindres carrés

Soit S une série statistique à deux variables quantitatives discrètes X et Y de taille $n \in \mathbf{N}^*$ définie par $S = \{(x_i\ ; y_i)\}_{1 \leq i \leq n}$. On appelle droite d'ajustement affine par la méthode des moindres carrés de S la droite d'équation $y = \frac{\text{cov}(X, Y)}{V(X)}\ (x - x) + \overline{y}$ où cov (X,Y) est la covariance de X et Y, V(X) la variance de X, et $\overline{x}$ et $\overline{y}$ les moyennes arithmétiques respectives de X et Y.

Écart interquartile

Soit S une série statistique à une variable quantitative discrète. L'écart interquartile de S est la longueur de l'intervalle interquartile.

Écart type

Soit *a* une série statistique de variance V. On appelle écart type *s* la racine carrée de la variance soit $s = \sqrt{V}$.

Effectifs cumulés croissants

Soit S une série statistique à une variable de type quantitatif et *a* une modalité de S. L'effectif cumulé croissant associé à *a* est la somme des effectifs de toutes les modalités inférieures ou égales à *a* dans la série S, c'est-à-dire le nombre d'individus de la population pour lesquels le caractère étudié a une modalité inférieure ou égale à *a*.
Dans le cas d'une série S dont les modalités sont regroupées par classes, l'effectif cumulé croissant de la classe $[a\ ; b[$ est la somme des effectifs de cette classe et des classes qui précèdent (c'est-à-dire dont les éléments sont strictement inférieurs à *a*) s'il y en a, autrement dit le nombre d'individus de la population pour lesquels le caractère étudié a une modalité appartenant à cette classe ou à l'une des classes qui précèdent s'il y en a.

Effectifs cumulés décroissants

Soit S une série statistique à une variable de type quantitatif et *a* une modalité de S. L'effectif cumulé décroissant associé à *a* est la somme des effectifs de toutes les modalités supérieures ou égales à *a* dans la série S, c'est-à-dire le nombre d'individus de la population pour lesquels le caractère étudié a une modalité supérieure ou égale à *a*.
Dans le cas d'une série S dont les modalités sont regroupées par classes, l'effectif cumulé croissant de la classe $[a\ ; b[$ est la somme des effectifs de cette classe et des classes qui suivent (c'est à dire dont les éléments sont strictement supérieurs ou égaux à *b*) s'il y en a, autrement dit le nombre d'individus de la population pour lesquels le caractère étudié a une modalité appartenant à cette classe ou à l'une des classes qui suivent s'il y en a.

Équation réduite d'une droite

Le plan est muni d'un repère et soit D une droite du plan.
– Si D n'est pas parallèle à l'axe des ordonnées, alors il existe un unique couple de réels (m, p) tel que l'équation $y = mx + p$ soit une équation de D. Cette équation est appelée équation réduite de D.
– Si D est parallèle à l'axe des ordonnées, alors il existe un unique réel *c* tel que l'équation $x = c$ soit une équation de D.

Étendue

Soit *a* une série statistique quantitative discrète à une variable. On appelle étendue de *a* la différence entre la plus grande et la plus petite valeur des modalités de *a*.

Événement

Soit Ω l'univers associé à une expérience aléatoire. On appelle événement toute partie de Ω. Un événement élémentaire est un événement qui ne contient qu'une seule issue.

Événements contraires

Soit Ω l'univers associé à une expérience aléatoire et *A* un événement de cet univers. L'événement contraire de *A*, noté $\overline{A}$, est l'ensemble de toutes les issues de l'univers Ω n'appartenant pas à *A*.

Événement impossible

Soit Ω l'univers associé à une expérience aléatoire et *A* un événement de cet univers. L'événement A est appelé impossible, noté $\emptyset$, lorsque qu'aucune éventualité de Ω vérifie *A*.

Événements incompatibles

Soit *A* et *B* deux événements d'un même univers Ω. *A* et *B* sont deux événements incompatibles si et seulement si l'événement $A \cap B$ est l'événement impossible. Cela veut dire qu'aucune issue ne peut réaliser à la fois *A* et *B*, ou encore que la réalisation de l'un des deux événements *A* et *B* exclut celle de l'autre.

Événements indépendants

Soit *A* et *B* deux événements d'un même univers Ω. Si *A* est de probabilité non nulle, *A* et *B* sont indépendants si et seulement si $p_A(B) = p(B)$. De même, si *B* est de probabilité non nulle, *A* et *B* sont indépendants si et seulement si $p_B(A) = p(A)$.

Fonction affine

On appelle fonction affine toute fonction *f* pour laquelle il existe deux réels *a* et *b* tels que pour tout réel *x*, $f(x)= ax + b$.

Fonction dérivée

Soit *f* une fonction définie sur un intervalle I, *f* est dite dérivable sur I si et seulement si, pour tout $a \in I$, *f* admet un nombre dérivé en *a*. La fonction définie sur I qui, à tout réel *a* de I, associe le nombre dérivé de *f* en *a* est appelée fonction dérivée de *f*. Cette fonction est notée *f* '.

Fonction linéaire

On appelle fonction linéaire toute fonction *f* pour laquelle il existe un réel *a* tel que pour tout réel *x*, $f(x) = ax$.

Forme canonique

Soit $a \in \mathbf{R}^*$, *b* Œ **R**, $c \in \mathbf{R}$, et *f* la fonction du second degré définie pour tout $x \in \mathbf{R}$ par $f(x) = ax^2 + bx + c$. Pour tout $x \in \mathbf{R}$, $f(x) = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a}$ est la forme

canonique de *f*.

Fréquence

Soit S une série statistique à une variable et *a* une modalité de S. La fréquence associée à *a* est le quotient de l'effectif de cette modalité par l'effectif total de la série S.
Dans le cas d'une série S de type quantitatif dont les modalités sont regroupées en classes, la fréquence d'une classe est le quotient de l'effectif de cette classe par l'effectif total de la série S.

Fréquences cumulées croissantes

Soit S une série statistique à une variable de type quantitatif et *a* une modalité de S. La fréquence cumulée croissante associée à *a* est la somme des fréquences de toutes les modalités inférieures ou égales à *a* dans la série S.
Dans le cas d'une série S dont les modalités sont regroupées en classes, la fréquence cumulée croissante de la classe $[a ; b[$ est la somme des fréquences de cette classe et des classes qui précèdent (c'est-à-dire dont les éléments sont strictement inférieurs à *a*) s'il y en a.

Fréquences cumulées décroissantes

Soit S une série statistique à une variable de type quantitatif et *a* une modalité de S. La fréquence cumulée décroissante associée à *a* est la somme des fréquences de toutes les modalités supérieures ou égales à *a* dans la série S.
Dans le cas d'une série S dont les modalités sont regroupées en classes, la fréquence cumulée croissante de la classe $[a; b[$ est la somme des fréquences de cette classe et des classes qui suivent (c'est-à-dire dont les éléments sont supérieurs ou égaux à *a*) s'il y en a.

Histogramme

Un histogramme est une représentation graphique d'une série statistique de variable quantitative. Il est constitué de rectangles contigus dont les aires sont proportionnelles aux effectifs de chaque classe. Sur l'axe des abscisses sont reportées les bornes des classes de la série. La suite des aires de ces rectangles est proportionnelle à la suite des effectifs des classes qu'ils représentent.

Image d'un nombre réel par une fonction numérique

Soit I un intervalle de **R**, f une fonction de I dans **R** et x un réel de I.
La fonction f associe à x un unique réel appelé image de x par la fonction f, noté $f(x)$.

Intervalle interquartile

Soit S une série statistique quantitative discrète à une variable. L'intervalle interquartile de S est l'intervalle dont les bornes sont le premier et le troisième quartile.

Inverse

L'inverse d'un nombre réel non nul a est le nombre b tel que $a \times b = 1$.

Issue

Une expérience aléatoire est une expérience dont on ne peut pas prévoir avec certitude le résultat. On appelle issue d'une expérience aléatoire tout résultat possible de cette expérience.

Maximum

Soit f une fonction définie sur un intervalle I. f admet un **maximum** sur I, lorsqu'il existe un nombre réel c appartenant à I et tel que, pour tout x de I, $f(x) \leq f(c)$.
Le nombre $M = f(c)$ est le **maximum** de f sur I.

Minimum

Soit f une fonction définie sur un intervalle I. f admet un **minimum** sur I, lorsqu'il existe un nombre réel c appartenant à I et tel que, pour tout x de I, $f(x) \geq f(c)$.
Le nombre $M = f(c)$ est le **minimum** de f sur I.

Médiane

Soit a une série statistique quantitative discrète à une variable de taille $n \in \mathbf{N}^*$ définie par : $a = \{a_i\}_{1 \leq i \leq n}$ ordonnée dans l'ordre croissant.
– Si n est impair avec $n = 2p + 1$ ($p \in \mathbf{N}^*$), alors une médiane de a est égale à a_{p+1} ;
– Si n est pair avec $n = 2p$ ($p \in \mathbf{N}^*$) alors une médiane de a est égale à $\dfrac{a_p + a_{p+1}}{2}$.

Mode

Soit a une série statistique quantitative discrète à une variable. Un mode est une modalité de a de plus grand effectif.

Moyenne arithmétique

Soit a une série statistique quantitative discrète à une variable de taille $n \in \mathbb{N}^*$, définie par $\{a_i\}_{1 \leq i \leq n}$.
La moyenne arithmétique de a, notée $\overline{a}$, est donnée par $\overline{a} = \dfrac{1}{n}(a_1 + a_2 + \dots + a_n)$.

Dans le cas où chaque modalité a_i apparaît avec l'effectif n_i, la moyenne arithmétique $\overline{a}$ de a peut s'écrire $\overline{a} = \dfrac{1}{N}(n_1 a_1 + n_2 a_2 + \dots + n_p a_p)$ où $N = n_1 + n_2 + \dots + n_p$.

Nuage de points

Soit S une série statistique à deux variables quantitatives discrètes X et Y de taille $n \in \mathbf{N}^*$ définie par $S = \{(x_i ; y_i)\}_{1 \leq i \leq n}$. On appelle nuage de points associé à S l'ensemble des points M_i de coordonnées $(x_i ; y_i)$ où $1 \leq i \leq n$.

Opposé

L'opposé d'un nombre réel a est le nombre b tel que $a + b = 0$.

Ordonnée à l'origine d'une droite

Le plan est muni d'un repère. Soit D une droite du plan non parallèle à l'axe des ordonnées. Il existe un unique couple (m, p) tel que la droite D ait pour équation réduite $y = mx + p$. Le réel p est appelé ordonnée à l'origine de la droite D.

Point moyen d'un nuage

Soit S une série statistique à deux variables quantitatives discrètes X et Y de taille $n \in \mathbf{N}^*$ définie par $S = \{(x_i ; y_i)\}_{1 \leq i \leq n}$. On note $\overline{x}$ et $\overline{y}$ les moyennes arithmétiques de $X = \{x_i\}_{1 \leq i \leq n}$ et de $Y = \{y_i\}_{1 \leq i \text{ £ } n}$. Soit $\{M_i\}_{1 \leq i \leq n}$ le nuage de points associé à S où M_i est le point de coordonnées $(x_i ; y_i)_{1 \leq i \leq n}$. Le point moyen du nuage $\{M_i\}_{1 \leq i \leq n}$ est le point G de coordonnées $(\overline{x}, \overline{y})$.

Pourcentage

– Instantané : soit E un ensemble de référence dont l'effectif est n, entier naturel non nul et A un sous-ensemble de E d'effectif n_A, le quotient $q = \frac{n_A}{n}$ est appelé **pourcentage instantané** de A par rapport à E.
– D'évolution (ou taux d'évolution) : soit x un nombre réel positif variable prenant la valeur non nulle x_0 à la date t_0 et pour valeur x_1 à la date t_1.
Le nombre p défini par $p = 100 \times \frac{x_1 - x_0}{x_0}$ est appelé **pourcentage d'évolution** de x de la date t_0 à la date t_1. Il sera noté p %.

Proportion

– Quantiles : soit S une série statistique à une variable quantitative discrète de taille n ordonnée dans l'ordre croissant.
La fonction quantile Q est la fonction définie sur $\left[\frac{1}{n} ; 1\right]$ qui, à tout réel $x \in \left[\frac{1}{n} ; 1\right]$, associe le terme a_i de S dont l'indice i est le plus petit entier supérieur ou égal à nx.

- Les trois quartiles d'une série S comprenant au moins 4 termes sont donnés respectivement par : Q (0,25), Q (0,50), Q (0,75).
- Les neuf déciles d'une série S comprenant au moins 10 termes sont donnés par $Q\left(\frac{i}{10}\right)$ où $i \in \{1, 2\ldots, 9\}$.
- Les 99 centiles d'une série S comprenant au moins 100 termes sont donnés par $Q\left(\frac{i}{100}\right)$ où $i \in \{1, 2\ldots, 99\}$.

– Quartiles : soit S une série statistique à une variable quantitative discrète ordonnée dans l'ordre croissant.
Le premier quartile Q_1 de S est le plus petit élément a de S tel qu'au moins 25 % des données soient inférieures ou égales à a.
Le troisième quartile Q_3 de S est le plus petit élément $a¢$ de S tel qu'au moins 75 % des données soient inférieures ou égales à a.

Somme des termes consécutifs d'une suite arithmétique

Soit $(u_n)_{n \in \mathbf{N}}$ une suite arithmétique.
Pour tout $n \in \mathbf{N}$: $u_0 + u_1 + \ldots + u_n = (n + 1)\frac{u_0 + u_n}{2}$.
Pour tout $n \in \mathbf{N}$ et tout $p \in \mathbf{N}$ tels que $p \leq n$: $u_p + u_{p+1} + \ldots + u_n = (n - p + 1)\frac{u_p + u_n}{2}$
c'est-à-dire : *nombre de termes* $\times \frac{\textit{premier terme de la somme + dernier terme de la somme}}{2}$.

Somme des termes consécutifs d'une suite géométrique

Soit $(u_n)_{n \in \mathbf{N}}$ une suite géométrique de raison q différente de 1.

Pour tout $n \in \mathbf{N}$: $u_0 + u_1 + \ldots + u_n = u_0 \dfrac{1 - q^{n+1}}{1 - q}$.

Pour tout $n \in \mathbf{N}$ et tout $p \in \mathbf{N}$ tels que $p \leq n$: $u_p + u_{p+1} + \ldots + u_n = u_p \dfrac{1 - q^{n-p+1}}{1 - q}$

c'est-à-dire : premier terme de la somme × $\dfrac{1 - q^{\text{nombre de termes de la somme}}}{1 - q}$.

Suite arithmétique

Une suite $(u_n)_{n \in \mathbf{N}}$ est dite suite arithmétique si et seulement si, il existe un réel r tel que, pour tout $n \in \mathbf{N}$, on ait $u_{n+1} = u_n + r$. Le réel r est appelé **raison** de la suite.
Si $(u_n)_{n \in \mathbb{N}}$ est une suite arithmétique de raison r, alors :
– Pour tout $n \in \mathbf{N}$ et pour tout $p \in \mathbf{N}$ tel que $p \leq n$, on a $u_n = u_p + (n - p)r$.
– Pour tout $n \in \mathbf{N}$, $u_n = u_0 + nr$.

Suite géométrique

Une suite $(u_n)_{n \in \mathbf{N}}$ est dite suite géométrique si et seulement si il existe un réel q non nul tel que, pour tout $n \in \mathbf{N}$, on ait $u_{n+1} = q\, u_n$. Le réel non nul q est appelé **raison** de la suite.
Si $(u_n)_{n \in \mathbf{N}}$ est une suite géométrique de raison q, alors :
– Pour tout $n \in \mathbf{N}$ et pour tout $p \in \mathbf{N}$ tel que $p \leq n$, on a $u_n = u_p \times q^{n-p}$.
– Pour tout $n \in \mathbf{N}$, $u_n = u_0 \times q^n$.

Tangente à une courbe

Soit f une fonction numérique définie sur un intervalle I. Soit a un réel de I tel que f soit dérivable en a, de nombre dérivé $\mathrm{L} = f'(a)$. On note C la courbe représentative de f dans un repère orthogonal du plan. La tangente à C au point A $(a\,;f(a))$ est la droite passant par A et de coefficient directeur L.

Taux d'accroissement

Soit a et b deux nombres réels avec a non nul. Le taux d'accroissement entre a et b est le nombre t défini par : $t = \dfrac{b - a}{a}$.

Taux de variation

Soit f une fonction numérique définie sur un intervalle I, et a, b deux réels distincts de I. Le taux de variation de f entre a et b est le nombre réel m défini par : $m = \dfrac{f(b) - f(a)}{b - a}$.

m est le coefficient directeur de la droite passant par les points A et B de coordonnées respectives $(a\,; f(a))$ et $(b\,; f(b))$.

Troncature

Soit x un nombre décimal positif.
– La troncature de x à l'unité est le nombre entier t tel que : $t \leq x < t + 1$.
– La troncature de x au dixième (ou à une décimale) est le nombre décimal t tel que $10t$ est un entier et : $t \leq x < t + 0{,}1$.
– La troncature de x au centième (ou à deux décimales) est le nombre décimal t tel que $100t$ est un entier et : $t \leq x < t + 0{,}01$.
– La troncature de x au millième (ou à trois décimales) est le nombre décimal t tel que $1000t$ est un entier et : $t \leq x < t + 0{,}001$.
Soit x un nombre décimal positif et n un nombre entier naturel. La troncature de x à 10^{-n} est le nombre tel que $10^n \times t$ est un entier et $t \leq x < t + 10^{-n}$.

Valeur approchée

Soit x et a deux nombres réels et p un nombre réel strictement positif. Le nombre a est une valeur approchée de x à p près, ou à la précision p, si et seulement si $|x - a| \leq p$ c'est-à-dire si et seulement si $a - p \leq x \leq a + p$.

Valeur approchée par défaut

Soit x et a deux nombres réels, et p un nombre réel strictement positif. Le nombre a est une valeur approchée par défaut de x à p près, ou à la précision p, si et seulement si a est une valeur approchée de x à p près inférieure à x, c'est-à-dire si et seulement si $a \leq x \leq a + p$.

Valeur approchée par excès

Soit x et a deux nombres réels, et p un nombre réel strictement positif. Le nombre a est une valeur approchée par excès de x à p près, ou à la précision p, si et seulement si a est une valeur approchée de x à p près supérieure à x, c'est-à-dire si et seulement si $a - p \leq x \leq a$.

Variance

On appelle variance V la moyenne des carrés des écarts à la moyenne
$\mathrm{V} = \frac{1}{\mathrm{N}} \sum_{i=1}^{n} n_i (x_i - \overline{x})^2$.

Références

Bibliographie

– John W. Tukey Addison-Wesley, *Exploratory Data Analysis*, Publishing Company.
– Marie-Christine Clugnet, *Calcul de doses*, MED-LINE.
– P. Dutarte et J.-L. Piednoir, *Enseigner la statistique au lycée : des enjeux aux méthodes*, Commission Inter-IREM, lycées technologiques IREM Paris-Nord.
– Arthur ENGEL, *Les Certitudes du hasard*, Aléas.
– B. Falissard, *Comprendre et utiliser les statistiques dans les sciences de la vie*, Masson.
– Bernard Legras, *Éléments de statistique à l'usage des étudiants en médecine et en biologie – cours et exercices corrigés*, Ellipses.
– Sabin Lessard-Monga, *Statistiques, concepts et méthodes*, Masson.
– Martine Peguin, *Calcul de doses et soins infirmiers*, Éditions Lamara.
– Dominique Rispail et Alain Viaux, *Guide du calcul de doses et de débits médicamentaux*, Masson.
– Daniel Schwartz, *La Statistique et le Vivant*, Flammarion.
– Daniel Schwartz, *Méthodes statistiques à l'usage des médecins et des biologistes*, Flammarion.
– P. Trousset et J.-F. Morin, *Mathématiques pour les sciences de la Terre,* tome 2, *Probabilités et statistique*, Édisciences International.
– *Biostatistique et probabilités – exercices, problèmes et épreuves corrigées*, Ellipses, coll. « PCEM ».
– *Dé-chiffrer par les maths, Pour un regard critique sur le monde*, APMEP / IREM de Lorraine, brochure APMEP, n° 147.
– *Enseigner les probabilités au lycée*, Commission Inter-IREM statistiques et probabilités, IREM de Reims.
– « Les mathématiques sociales », *Pour la science*, hors-série n° 24, juillet 1999.
– Publications de l'association « Pénombre ».
– « Spécial probabilités », *Tangente*, n° 47, décembre 1995.
– « Spécial probabilités », *Repère IREM*, n° 32, juillet 1998.

Sitographie

– Banque de données en santé publique (BDSP) – Institut de recherche et de documentation en économie de la santé (IRDES, ex-CREDES – Centre de recherche et de documentation en économie de la santé) : www.irdes.fr
– Comité français d'éducation pour la santé : www.cfes.santé.fr
– Documentation française (données démographiques) : www.ladocumentationfrancaise.fr
– Haut comité en santé publique (HCSP) : www.hcsp.ensp.fr
– Institut national d'études démographiques (INED) : www.ined.fr
– Institut national de la santé et de la recherche médicale (INSERM) : www.inserm.fr
– Institut national de la statistique et des études économiques (INSEE) : www.insee.fr
– Medcost, Internet médical, santé et systèmes d'informations médicaux : www.medcost.fr
– Ministère de la Santé : www.sante.gouv.fr

– Réseau national de ressources en sciences médico-sociales (RNRSMS) : www.ac-creteil.fr/sms
– Statistiques médicales en ligne : www.imc.imag.fr/Imc-sms/Bernard.Ycart/smel/
– informations et sites consacrés à Florence Nightingale :
• www.ibe.unesco.org/publications/ThinkersPdf/nightinf.PDF#search=%22%20florence%20nightingale%22
• www.1914-1918.be/soigner_florence_nightingale.php
• www.florence-nightingale.co.uk/
• www-history.mcs.st-andrews.ac.uk/history/Mathematicians/Nightingale.html
– www4.ac-lille.fr/~math/classes/stats/piednoir/intervention2.html
– Site Statistix Euler : www.statistix.fr
– Site Musibiol, sujets de biologie donnés au baccalauréat : http://musibiol.net/biologie/