

Document d'accompagnement de 1^{ère} ES

GEPS

version du 30/01/01

Le programme publié dans le BO du 31 août 2000 (hors-série n°8) expose de façon détaillée les objectifs de l'enseignement des mathématiques en classe de 1^{ère} ES ; il convient de le relire attentivement avant tout travail à l'aide du présent document. Le paragraphe 4 du programme détaille les domaines mathématiques à traiter en première ES : les deux premières colonnes (contenus et modalités de mise en œuvre) s'imposent à chacun, la troisième fournit quelques commentaires que le présent document vient compléter.

Ce document d'accompagnement évoluera au cours de l'année en fonction des contributions et demandes des divers intervenants du système scolaire ; la version définitive pour l'année 2001-02 sera fournie au printemps 2001.

1. Généralités

Comme indiqué en introduction du programme, il a été choisi de proposer un enseignement de mathématiques consistant et adapté à la spécificité de la série ES.

On a donc privilégié quelques outils indispensables au traitement mathématique de l'information chiffrée ; d'où les deux grands titres fédérant les diverses notions abordées dans la partie obligatoire du programme : traitement des données et probabilités d'une part, algèbre et analyse d'autre part.

Pour l'option, les thèmes abordés sont relativement indépendants de la partie obligatoire ; ils sont par contre partiellement corrélés à ceux de la spécialité de terminale ES : l'élève désirant choisir la spécialité mathématique en terminale aura donc intérêt à choisir l'option en première (une telle position est cohérente avec le souci de responsabiliser les élèves dans leur choix de formation : la première n'est plus une classe d'orientation et il est normal que les choix faits pour la classe de première aient une incidence jusqu'en terminale).

Un apprentissage actif

Le passage à l'abstraction mathématique présente parfois plus de difficultés pour les élèves de cette section : il importe donc de veiller au caractère progressif et actif de l'apprentissage proposé à chaque élève. Des exemples seront systématiquement utilisés pour introduire les nouveaux concepts. Ceux-ci seront ensuite mis en œuvre dans des exercices et problèmes qui en montreront la portée.

Expression et raisonnement

Les résultats acquis (théorèmes, techniques opératoires,...) seront clairement énoncés et leur statut mathématique explicité. On veillera à la qualité des raisonnements effectués ; dans le champ mathématique, les modalités du raisonnement sont variées ; la plupart ne sont pas propres aux mathématiques : formulation des questions et argumentation logique se retrouvent ainsi dans la dissertation en français, philosophie, lors de travaux de sciences expérimentales... Par ailleurs dans toutes les disciplines, c'est la même langue que l'on utilise, avec ses règles syntaxiques, ses mots de liaison logique ; à ce titre, le professeur de mathématiques participe au travail de maîtrise de la langue française, de la syntaxe tout particulièrement, et il importe que les élèves en aient conscience.

Importance de l'acquisition de quelques automatismes libérateurs.

Le calcul est une activité qui dans sa diversité (calcul numérique, algébrique, vectoriel, etc.) est assez spécifique des mathématiques ; le calcul sous diverses formes est omniprésent, sa pratique régulière est indispensable pour avancer en mathématiques. Calculs simples et règles opératoires mués en automatismes libèrent la pensée, facilitent la compréhension et permettent de se consacrer à d'autres tâches. La pratique régulière et individuelle des calculs que l'on peut faire à la main, est une part indispensable et un moyen irremplaçable d'accéder à une bonne compréhension. Il ne s'agit pas de rechercher la virtuosité : calculatrices et ordinateurs ont rendu cet objectif dérisoire. Il ne suffit pas pour autant d'obtenir des résultats sur une machine : il faut d'abord anticiper un calcul, au moins dans sa forme, pour percevoir l'intérêt de sa mise en œuvre, puis savoir interpréter les résultats, juger de leur validité et de leur limite.

L'aide des outils de calcul

La mise en œuvre anticipée du programme de seconde dans une vingtaine de lycées durant l'année 1999-2000 a confirmé le grand intérêt de l'usage d'outils logiciels (ordinateurs et calculatrices) pour faciliter l'approche et l'assimilation de certains concepts. D'où, entre autres, les invitations nombreuses du programme à utiliser ces outils : l'expérience et la généralisation d'une telle pratique devraient permettre d'en tirer tout le bénéfice attendu.

2. A propos du titre "Traitement des données et probabilités"

Le choix a été fait pour la section ES de donner un rôle important aux séries chronologiques, particulièrement fréquentes dans les cours d'économie de cette section ; mais on veillera à ce que les questions et exemples traités conduisent jusqu'à une réflexion conceptuelle ou axiomatisée, ce qui constitue une bonne préparation à d'éventuelles études ultérieures davantage centrées sur des pratiques professionnelles.

a) Pourcentages

Comme indiqué dans le programme, il ne s'agit pas d'aborder **ici** quelque connaissance technique nouvelle, mais d'entretenir un apprentissage de base indispensable pour lire correctement et de façon critique l'information chiffrée. Ce paragraphe recoupe de nombreux autres titres du programme : statistique (fréquences, données en pourcentage, évolution de séries chronologiques,...), suites géométriques (obtenues par augmentations successives), dérivation (approximation affine), ... Les notions qu'il décrit devront donc être régulièrement mises en jeu, en particulier à partir de données issues des médias ou de l'environnement scolaire de la classe.

b) Nature des données

On a parfois prôné, pour l'enseignement de la statistique, le recueil de données par les élèves eux-mêmes, cette pratique étant considérée comme motivante et permettant de percevoir le champ de l'aléatoire. Or la perception de l'aléatoire peut s'acquérir de manière plus profonde par la simulation. Par ailleurs, pour de nombreuses questions que l'on peut se poser dans le champ scolaire, des données existent : elles sont réactualisées chaque année, leur contenu est riche et elles sont accessibles dans des banques de données. Sans exclure complètement un recueil ponctuel de données par les élèves à condition qu'il ne prenne pas beaucoup de temps, on s'appuiera avant tout sur des données existantes. Certaines données sont des données brutes (exemple : série des hauteurs d'un fleuve mesurées en un point géographique précis tous les jours à la même heure) ; d'autres sont obtenues en prenant des moyennes de mesures brutes (séries des températures mensuelles en un point géographique précis, évolution sur une période de 10 ans des dépenses de logement suivant les catégories socioprofessionnelles) ; certaines encore sont des moyennes mobiles : par exemple, dans la présentation des mesures de température quotidienne à une heure donnée en un point donné, on remplace parfois les températures brutes $x_0, \dots, x_n$ par la

série des moyennes mobiles d'ordre k , $y_k, \dots, y_{n-k}$, calculées ainsi : $y_i = \frac{1}{2k+1} \sum_{j=i-k}^{i+k} x_j$, pour $i=k, \dots, n-k$.

On pourra montrer sur des exemples, avec $k=2$ ou 3 , que la courbe obtenue en remplaçant les données brutes par les moyennes mobiles est plus lisse.

c) Effets de structure

Exemple : le revenu moyen global des individus actifs d'une population (par exemple la population parisienne) peut augmenter avec le temps alors que dans toutes les catégories socioprofessionnelles (CSP) le revenu baisse, l'augmentation globale étant liée à un changement de la répartition en CSP (à Paris, les CSP à faible revenu ont eu tendance à déménager en banlieue). La structure à une date donnée est ici la répartition des CSP à cette date.

d) Diagrammes en boîtes

Il ne s'agit pas là d'un élément à part ; il sera introduit à l'occasion du traitement de données expérimentales ou d'activités de simulation. On trouvera en annexe de ce document une note sur ce type de diagrammes à l'usage des enseignants.

e) Etude fréquentielle de tableaux à double entrée

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 ne construira pas les

« tableaux théoriques » (on n'introduira pas de façon formelle la notion de sur- et sous-représentation, celles-ci n'ayant vraiment de sens que si elles sont « significatives » au sens statistique).

Exemple : En 1979, le New-York Times a noté qu'entre 1973 et 1979, en Floride, il a été prononcé 131 peines capitales pour meurtre ; parmi ces condamnés, 55% étaient des Blancs alors que 48 % de ceux qui ont été jugés pour meurtre pendant la même période étaient des Blancs. Qu'en penser ?

Avant d'en penser quoique ce soit, il est opportun de s'intéresser aussi à la couleur de la victime.

1- Entre 1973 et 1979, il y a eu 2433 meurtres en Floride, dont la victime était de couleur blanche. Le tableau ci-dessous donne les détails suivant la couleur de la peau du suspect (blanche, B, ou non blanche, NB), les colonnes indiquant la sentence (PC : peine capitale ; AS : autre sentence)

	PC	AS	Totaux
NB	48	239	287
B	72	2074	2146
Totaux	120	2313	2433

- Dire comment est construit le tableau ci-dessous et ce que signifient les trois lignes :

	PC	AS	Totaux
NB	16,7	83,3	100,0
B	3,4	96,6	100,0
Totaux	4,9	95,1	100,0

Remarque : les différences de pourcentages de condamnation à la peine capitale suivant la couleur du suspect, lorsque la victime est de couleur blanche, est trop grande pour être imputée à la fluctuation d'échantillonnage.

2- Entre 1973 et 1979, il y a eu 4764 meurtres en Floride. Le tableau ci-dessous donne les détails suivant la couleur de peau du suspect (blanche, B, ou non blanche, NB), les colonnes indiquant la sentence (PC : peine capitale ; AS : autre sentence).

	PC	AS	Totaux
NB	59	2448	2507
B	72	2185	2257
Totaux	131	4633	4764

- Construire un tableau analogue au deuxième tableau de la question 1 et le commenter.

- Quel est le pourcentage de suspects de couleur blanche lorsque la victime est de couleur blanche ; quel est le pourcentage de suspects de couleur blanche lorsque la victime n'est pas de couleur blanche ?

- Faire l'étude lorsque la victime n'est pas de couleur blanche.

Remarque : les différences de pourcentages de condamnation à la peine capitale suivant la couleur du suspect, sans tenir compte de la couleur de la victime, ou lorsque la victime est non blanche, peuvent être imputées à la fluctuation d'échantillonnage, i.e. ne peuvent pas être interprétées comme une tendance des jugements en Floride.

Exemple : Que signifient les tableaux ci-dessous et comment sont-ils déduits du premier tableau ?

Les données concernent la répartition suivant le sexe et le poste (AG : agent spécial , P : personnel autre) des emplois au FBI au 31 janvier 1997.

	H	F			H	F		
AS	9199	1617	10816		AS	36,1	6,3	42,4
P	4535	10165	14700		P	17,8	39,8	57,6
	13734	11782	25516			53,8	46,2	100,0

	H	F			H	F	
AS	85,0	15,0	100,0	AS	67,0	13,7	42,4
P	30,9	69,1	100,0	P	33,0	86,3	57,6
	53,8	46,2	100,0		100,0	100,0	100,0

Remarque : on distinguera les calculs de l'interprétation, dans le contexte, des données d'observation.

f) Loi de probabilité

On recensera les propriétés mathématiques élémentaires de l'objet « distributions de fréquences » (cf. tableau ci-dessous) et on définira une loi de probabilité comme un objet mathématique ayant les mêmes propriétés.

Distribution de fréquences sur $E=\{x_1, \dots, x_r\}$	Loi de probabilité sur $E=\{x_1, \dots, x_r\}$
$(f_1, \dots, f_r)$ $f_i \geq 0 ; \sum f_i = 1$	$(p_1, \dots, p_r)$ $p_i \geq 0 ; \sum p_i = 1$
$A \subset E$ Fréquence de A : $f(A) = \sum f_i$	$A \subset E$ probabilité de A : $P(A) = \sum p_i$
Événement complémentaire : $f(\bar{A}) = 1 - f(A)$	Événement complémentaire : $P(\bar{A}) = 1 - P(A)$
<i>Cas numérique :</i> <i>Moyenne empirique :</i> $\bar{X} = \sum f_i x_i$	<i>Cas numérique :</i> <i>Espérance d'une loi P :</i> $\mu = \sum p_i x_i$

Modéliser une expérience aléatoire, c'est lui associer une loi de probabilité (qui est un objet du monde mathématique).

Une fréquence est empirique : elle est calculée à partir de données expérimentales, alors que la probabilité d'un événement est un « nombre théorique » (un objet du monde mathématique). Les distributions de fréquences issues de la répétition d'expériences identiques et indépendantes varient (fluctuent), la loi de probabilité est un invariant associé à l'expérience.

L'objectif est que les élèves comprennent à l'aide d'exemples (cf. paragraphe sur les expériences de référence) que modéliser, c'est ici choisir une loi de probabilité. Il ne s'agit cependant en aucun cas de tenir des discours généraux sur les modèles et la modélisation.

Les élèves devront bien distinguer ce qui est empirique (du domaine de l'expérience) de ce qui est théorique ; en particulier, on réservera la lettre grecque σ à l'écart-type d'une loi et on évitera de noter avec cette même lettre un écart-type empirique (il s'agit là de règles de notation internationales).

En classe de première, une loi de probabilité P sur un ensemble fini est la liste des probabilités des éléments de E ; à partir de cette liste, on définit naturellement la probabilité d'événements, c'est-à-dire implicitement une application de $\mathcal{V}(E)$ dans $[0,1]$, application qui sera encore désignée par P.

Il est inutilement complexe, pour le cas des ensembles finis, de partir d'une application de $\mathcal{V}(E)$ dans $[0,1]$, vérifiant certains axiomes, puis de montrer ensuite que cette application est entièrement caractérisée par $(p_1, \dots, p_r)$. Le fait de ne pouvoir simplement généraliser cette définition aux ensembles continus, et la nécessité d'une définition ensembliste sera abordée en terminale.

Si tous les éléments d'un ensemble bien défini E ont même probabilité, celle-ci est dite équirépartie ; on parlera aussi d'équiprobabilité et on dira que les éléments de l'ensemble E sont choisis au hasard (i.e. si on ne spécifie rien de plus, le choix au hasard est un choix avec équiprobabilité). Si la loi P n'est pas équirépartie, on parlera de choix d'éléments selon la loi P (ou éventuellement de choix au hasard selon la loi P).

Dans le cas de choix au hasard, la probabilité d'un événement est le quotient de son nombre d'éléments par le nombre d'éléments de l'ensemble.

On évitera tout développement théorique sur le langage des événements et le calcul ensembliste qui en découle : ces notions et la pratique de la logique qu'ils impliquent (étude du complémentaire de l'événement "A ou B", ou de l'événement "A et B") s'acquièrent au fil d'exercices.

g) Modélisation d'expériences de référence

Modéliser une expérience aléatoire à valeurs dans un espace Ω , c'est choisir une loi de probabilité P définie sur Ω . Ce choix est en général délicat à faire, sauf dans certains cas où des considérations propres au protocole expérimental conduisent à proposer a priori un modèle. Il en est ainsi des lancers de pièces ou de dés pour lesquels des considérations de symétrie conduisent au choix d'un modèle où la loi de probabilité est équirépartie.

Sans faire une liste de conventions terminologiques, on indiquera clairement que les termes *équilibré* et *hasard* indiquent un choix du modèle de l'expérience, modèle où intervient quelque part une probabilité équirépartie.

h) Modélisation à partir de fréquences.

En dehors des cas évoqués ci-dessus où des considérations quant à la nature des expériences permettent de proposer un modèle, le choix d'un modèle à partir de données expérimentales est beaucoup plus délicat. On se contentera, pour certains exercices, de fournir un modèle en indiquant dans un premier temps que des techniques statistiques ont permis de déterminer et de valider un tel modèle à partir de nombreuses données expérimentales.

Pour déterminer et/ou valider un modèle probabiliste, le premier outil dont on dispose est un théorème de mathématiques appelé loi des grands nombres, dont un énoncé intuitif est :

Dans le monde théorique défini par une loi P sur un ensemble Ω , les fréquences des éléments de Ω dans une suite de n expériences identiques et indépendantes tendent vers leur probabilité quand n augmente indéfiniment.

On donnera des exemples où un modèle est déterminé à partir de fréquences ; les exemples les plus compliqués que l'on abordera consistent à associer une loi de probabilité à un tableau à double entrée.

i) Simulation

L'exemple ci-dessous montre comment mêler diverses composantes d'un travail mathématique : observation, premières conjectures, expérimentation à plus grande échelle, puis obtention et preuve de certains résultats :

Exemple : somme de deux dés.

En répétant 100 fois de suite le lancer de deux dés, on **observe** que certains résultats s'obtiennent plus souvent que d'autres.

À l'aide d'un tableur par exemple, il est possible d'**expérimenter** à plus grande échelle : simulation d'un plus grand nombre de lancers de deux dés et construction du tableau des effectifs : l'inégale répartition des fréquences de chaque résultat est flagrante

La recherche d'un modèle théorique adapté avec une loi de probabilité équirépartie permet ensuite calculs et démonstrations : on **prouve** que les résultats sont inégalement probables et on détermine précisément leur probabilité.

Résultats pour 5 séries de 100 lancers

2	0.02	0.02	0.01	0.03	0.01
3	0.03	0.02	0.09	0.06	0.04
4	0.08	0.08	0.08	0.08	0.1
5	0.11	0.11	0.12	0.09	0.11
6	0.17	0.13	0.1	0.11	0.15
7	0.15	0.15	0.12	0.14	0.2
8	0.12	0.2	0.1	0.13	0.22
9	0.15	0.1	0.12	0.13	0.06
10	0.09	0.1	0.14	0.1	0.08
11	0.06	0.05	0.06	0.09	0.06
12	0.02	0.04	0.06	0.04	0.01

3. A propos du titre "Algèbre et analyse"

*Le programme reprend pour l'essentiel le programme antérieur :
ce paragraphe ne sera donc pas davantage développé.*

Tableau des variations d'une fonction.

On admettra les conventions décrites dans l'exemple qui suit.

Supposons que l'étude des variations d'une fonction permette d'aboutir au tableau de variations suivant.

x	- ⑥	- 2	1	3	+ ⑥
f(x)	1	- 1	+ ⑥	5	-

On considérera que le tableau de variations est une forme stylisée de représentation graphique : le plus souvent, elle pourra suffire au niveau des productions des élèves, la calculatrice graphique donnant le produit fini. On admettra qu'il n'y a pas de « sauts » dans la courbe représentative sur un intervalle donné (notion intuitive de continuité qui sera précisée en terminale) et qu'une flèche inclinée correspond à une stricte monotonie. On pourra alors en déduire sans discours supplémentaire :

- ❑ des encadrements d'images,
- ❑ le nombre de zéros de f .

4. A propos de l'option

a) Compléments sur les fonctions

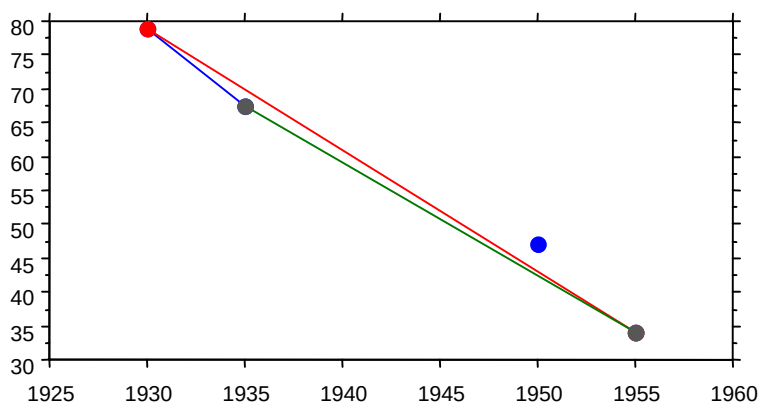
On recherchera des exemples de fonctions affines par morceaux dans les diverses disciplines de la section ES, mathématiques comprises ; on pourra envisager quelques cas de recollement de morceaux par continuité (tels le cas des fonctions impôt sur le revenu) : on s'appuiera pour ce faire sur l'intuition graphique et, sans le systématiser, on pourra rechercher l'expression explicite de $f(x)$ sur chaque "morceau". En liaison avec le paragraphe statistique, on pourra en particulier représenter la fonction $x \mapsto \frac{1}{n} \sum_{i=1}^n |x - x_i|$ dans le cas d'une série (x_i) de petite taille et vérifier que la moyenne ne minimise pas nécessairement cette fonction.

On ne traitera d'exemples d'interpolation linéaire que dans les cas où cette méthode a du sens.

Interpolation linéaire : exemple 1

Les taux de mortalité pour mille naissances en 1930, 1935, 1950, 1955 sont respectivement 78,8 67,6 47,2 34,2. Il manque les taux de mortalité en 1940 et 1945.

1. Au vu du nuage des 4 points correspondants aux données, un élève estime ces deux taux par interpolation linéaire entre 1935 et 1950, un autre par interpolation linéaire entre 1930 et 1955, un autre entre 1935 et 1955. Calculer ces résultats et les comparer.
2. Un quatrième élève estime que si on ne dispose pas des données en 1940 et en 1945, ce n'est pas fortuit et qu'en conséquence l'interpolation linéaire n'est peut-être pas pertinente...Qu'en pensez-vous ?
3. D'autres sources confirment les données pour 1930, 1935, 1950, 1955 et fournissent aussi les taux en 1940 et 1945 : 87,6 et 110. Pouvez vous donner un sens à la différence entre les valeurs calculées en 1 et les valeurs réellement observées ?



Interpolation linéaire : exemple 2

L'INED (www.ined.fr) fournit des données annuelles sur la population française. Les recensements n'étant organisés que tous les 6 à 9 ans (les derniers datent de 1999, 1990, 1982, 1975, 1968, 1962), détermine-t-on par interpolation linéaire les données pour les années hors recensement ?

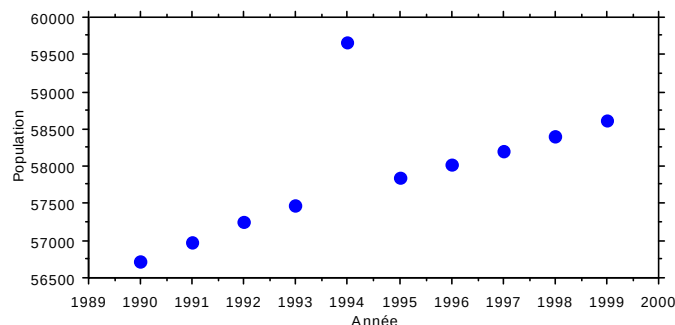
Voici pour information les "chiffres" de la population française (en milliers) fournis par l'INED pour les dix dernières années et recopiés :

1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
56 710	56 978	57 242	57 470	59 661	57 847	58 029	58 210	58 398	58 620

- La représentation graphique montre qu'il semble y avoir une erreur de recopie de données en 1994. Un élève suggère de remplacer cette donnée par la moyenne des années 1993 et 1995, un autre par la moyenne des données de 1994 et 1998, un autre par la moyenne des quatre années précédentes et des quatre années suivantes. Comparer les résultats.

- Comment interpréter le quasi-alignement des points du nuage, en dehors du point litigieux ?

Un élève retourne sur le site et lit, pour l'année 1994 une population de 57 661 ; il lit aussi sur le site que des recensements n'ont été faits qu'en 1990 et 1999 ; pour les autres années, on a appliqué les formules données par un modèle. Quels seraient les chiffres si comme modèle, on avait pris celui d'une croissance parfaitement linéaire entre 1990 et 1999 ?



b) Géométrie dans l'espace

C'est le seul endroit dans les programmes de première et terminale ES où les élèves ayant choisi l'option mathématiques travailleront sur la géométrie. Il s'agira avant tout, dans le fil de ce qui a été fait en classe de seconde, de géométrie "repérée" ; l'objectif est d'aider à une meilleure vision des objets de l'espace et de parvenir au traitement analytique de quelques problèmes. La difficulté liée à la représentation des objets de l'espace est grande : on évitera donc de multiplier les situations à traiter à la main et on s'appuiera autant que faire se peut sur des dessins fournis par des logiciels ou préparés par l'enseignant.

Pour traiter cette partie, on peut utiliser des logiciels de construction géométrique dans l'espace (tels GeospaceW) et des logiciels de calcul formel disponibles sur ordinateur ou sur calculatrice. Des exemples de séquences mettant en oeuvre ces logiciels peuvent être consultés sur le site Educnet, à partir de l'adresse :

<http://www.educnet.education.fr/math/reforme.htm>.

Exemple 1 : Elections au pays des Cartes

(Réécriture d'un énoncé de J. Lubczanski, dans "Les maths au jour le jour", éditions Cedic 1985)

Argine, Judith, Pallas et Rachel, sont les quatre prétendantes au titre de reine des reines.

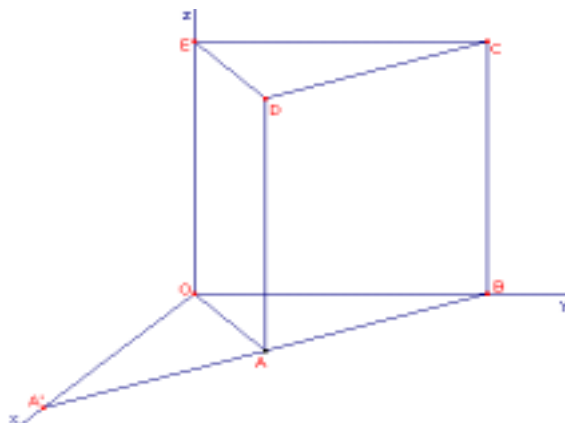
Au 1^{er} tour, Pallas obtient 19% des suffrages, Judith 33%, Rachel 16% et Argine 32%.

Seules, les deux candidates arrivées en tête peuvent se maintenir pour le second tour ; tous les prévisionnistes de la politique se plongent dans les calculs pour anticiper les résultats définitifs.

Ils admettent que :

- il n'y a pas de nouveaux électeurs au 2nd tour et les électeurs ayant voté au 1^{er} tour pour Argine ou Judith ne modifieront pas leur vote au 2nd tour ;
- parmi les électeurs ayant voté Pallas au premier tour, une proportion x reporteront leurs voix sur Judith au second tour, une proportion y , plus grande que x , se reporteront sur Argine et les autres s'abstiendront.
- parmi les électeurs ayant voté Rachel au premier tour, une proportion z voteront pour Judith au 2nd tour, et les autres s'abstiendront.

1) La donnée de (x,y,z) suffit donc pour déterminer le vote du second tour. Dans la représentation ci-dessous, on a visualisé graphiquement l'ensemble des situations de vote possibles au second tour (le repère $(O, \overrightarrow{OA'}, \overrightarrow{OB}, \overrightarrow{OE})$ est orthonormal) : décrire cet ensemble et contrôler le dessin.

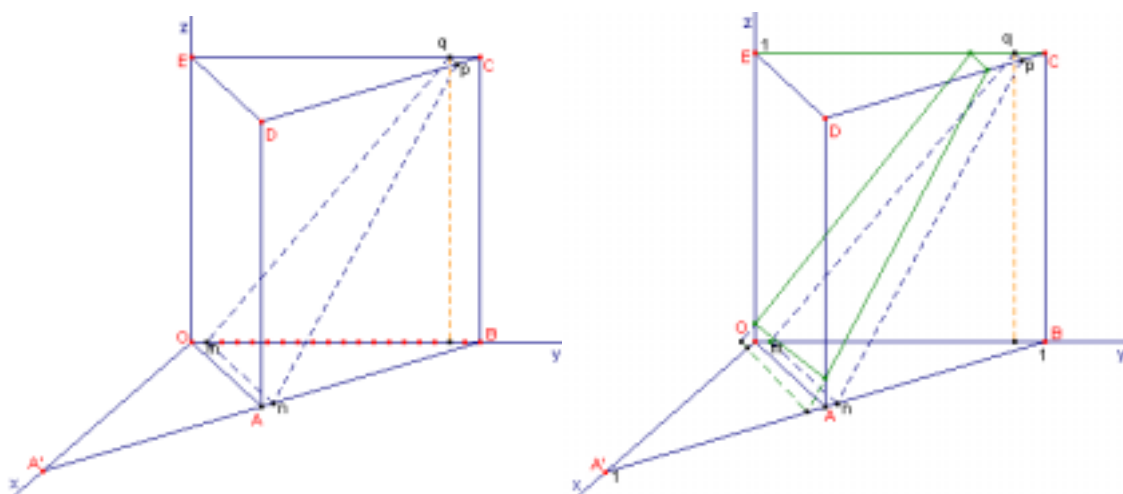


- 2) Quelles relations entre les trois variables x , y et z traduit le fait que Judith est élue au second tour ?
- 3) À quelles situations de vote correspondent les points A, D et C ?
- 4) À quelles régions de l'espace correspondent les situations suivantes : a) Judith est élue ? b) Judith et Argine sont ex aequo ? c) Argine est élue ?
- 5) Certains prévisionnistes assurent que :
 - a) si plus de 48% des électeurs de Pallas votent Judith, alors Judith est élue ;
 - b) si plus de 95% des électeurs de Pallas votent Argine, alors Argine est élue.

Leur prévisions vous semblent-elles correctes ? Quelle est la plus petite valeur m de x qui garantit l'élection de Judith ? Donner un encadrement des votes que peut avoir Argine lorsque $x=m$. Quelle est la plus petite valeur de y qui garantit l'élection d'Argine ?

6) Résultats

En fait, le résultat du second tour a donné : 51,5% pour Judith et 48,5% pour Argine (il s'agit des suffrages exprimés, c'est à dire que $A/J=48,5/51,5$). Trouver l'équation de l'ensemble des situations de vote correspondant à ce résultat.



Remarques : La recherche d'un tel problème permet la mise en œuvre de nombreux aspects de la géométrie dans l'espace (plans parallèles aux axes de coordonnées, équations de plans, équations de droites dans un plan, intersection de plan et de droites, etc.) ; à cette occasion, on admettra la propriété de régionnement de l'espace par un plan.

L'équation du plan des situations ex-aequo paraîtra peut-être compliquée à certains élèves ; on pourra éventuellement simplifier les calculs en considérant x , y et z comme des pourcentages de la population initiale (x représentant donc le pourcentage des suffrages exprimés au premier tour qui ont voté pour Pallas au premier tour et qui voteront pour Judith au second ;...).

Exemple 2 : Un problème de "programmation linéaire"

(en complément de ce qui est fait dans la partie obligatoire)

Un polygone des contraintes dans le plan Oxy

Une fonction coût du type $z = ax + by + c$

Dans l'espace,

le "cylindre" des solutions admissibles en (x,y) ;

le plan d'équation $z = ax + by + c$ et ses lignes de niveau représentés dans le plan Oxy ;

le point le plus haut de ce plan dans le "cylindre"

Exemple 3 : Courbes de niveau sur un cône, sur une sphère, sur un « ballon de rugby »

On considère un cône de révolution, une demi sphère et un ballon de rugby sur lesquels on a tracé des courbes de niveau.

Voici dans le désordre, comment ces courbes de niveau peuvent apparaître à un observateur placé au dessus.

On peut être étonné de la différence des dessins. À quel solide faut-il attribuer chacun des dessins ci-contre ?

Un logiciel de géométrie dynamique permet de les construire géométriquement sans effectuer de mesure.

Le calcul permet de préciser la suite des rayons obtenus dans le cas du cône et de la demi-sphère. Pour le ballon de rugby, le calcul n'est pas du niveau de 1^{ère}.

Cet exemple permet d'aborder, à partir d'objets géométriques élémentaires (cône, sphère), la notion de fonction de deux variables (ici $f(x,y) = x^2 + y^2$, ou $f(x,y) = \sqrt{r^2 - x^2 - y^2}$), de lire en contrôlant par le calcul quelques courbes de niveau, et de comprendre jusqu'à quel point on peut reconstituer un objet de l'espace à partir de coupes parallèles à un plan donné.

c) Calcul matriciel

L'objectif est ici la description puis la résolution bien comprise de problèmes faisant intervenir de nombreuses données numériques. Dans un premier temps, on montrera l'intérêt de tableaux de nombres pour présenter certaines informations chiffrées : tableaux à 1 ligne ou 1 colonne que l'on appellera vecteur-ligne ou vecteur-colonne, tableaux à n lignes et p colonnes que l'on appellera matrice d'ordre $n \times p$. On prendra du temps pour poser correctement l'interprétation matricielle de la situation à décrire ; on n'hésitera pas, dans cette phase initiale, à légender le tableau de nombres pour faciliter la compréhension.

Des exemples permettront ensuite d'introduire et d'illustrer les opérations (addition, multiplication par un nombre, multiplications) sur ces matrices (nous donnons quelques exemples ci-dessous). Il importe que les premiers calculs soient faits avec soin à la main ; on utilisera ultérieurement les outils de calcul (calculatrice, tableur, logiciel de calcul). Les premières propriétés de ces opérations sur les matrices (notamment associativité et non commutativité du produit de matrices carrées) seront mises en évidence (sans pourtant introduire le terme d'associativité).

Les matrices seront ensuite utilisées en complément du programme de la partie commune du programme de 1^{ère} ES pour résoudre des problèmes conduisant à des systèmes d'équations linéaires ; dans le cas où il y a unicité de la solution, cela amènera à la recherche de l'inverse d'une matrice, mais aucune technicité ne sera exigée en la matière.

Les matrices interviendront dans le programme de spécialité de terminale ES pour résoudre des problèmes de théorie des graphes.

Proposition de curriculum (fournie lors de la consultation du printemps 2000):

1. Vecteur-ligne (ou colonne), ou n-uplet.

Exemples :

- le portrait scolaire d'un élève en fin de premier trimestre : (10,7,9,12,15,10), où l'on a les notes de cet élève en français, mathématiques, économie, histoire-géographie, LV1 et EPS ;

- le travail d'une station-service durant une journée : (2500,500,1800,3000), donnant les quantités (en litres) d'essence SP95, essence SP98, super et gazoil ;
- le portrait politique d'un arrondissement : (15,10,3), donnant le nombre (en milliers) d'électeurs des partis A, B et C ;

- ...

D'où la définition d'un vecteur ou n-uplet de réels $\mathbf{A} = (a_1, a_2, \dots, a_n)$ (que l'on pourra aussi noter en colonne).

Dimension d'un vecteur ; égalité de deux vecteurs ; somme de deux vecteurs ; multiplication par un réel ; propriétés de ces opérations.

Exercices : le travail d'une station service durant une semaine ; le portrait politique d'un département, d'une région ; le portrait politique en % ;...

2. Matrices.

Exemples :

- notes de 5 élèves d'une classe ;
- travail des 4 stations service d'une agglomération ;
- répartition des élèves d'un lycée selon l'orientation et la CSP du père ;
- ...

Définition d'une matrice $n \times p$ à n lignes et p colonnes (ou tableau de nombres) ; notation (a_{ij}) .

Egalité de deux matrices ; matrice transposée ; un vecteur est une matrice $1 \times p$ ou $n \times 1$; matrice carrée, matrice carrée unité.

Addition de deux matrices de même dimension ; multiplication par un scalaire.

Exercices : répartition des élèves de 2 lycées d'une agglomération... ; expression en pourcentage du travail de quatre stations service...

3. Multiplication de matrices.

Exemple 1 : chiffre d'affaires de quatre stations service

Définition de AX où A est une matrice $n \times p$ et X un vecteur-colonne de dimension p (donc une matrice de dimension $p \times 1$).

Exemple 2 : matrice A , 4×3 , représentant le nombre de pièces de trois produits P_1 , P_2 et P_3 fabriqués par quatre artisans durant un an ; matrice B , 3×2 , donnant le nombre d'heures et le prix de la matière première pour fabriquer une pièce de chacun des produits P_1 , P_2 et P_3 . Comment obtenir le nombre d'heures et le prix de matière première par artisan par année ?

D'où la définition du produit de deux matrices.

Exercices :

- A^2 où A est une matrice de "connaissance" ou d'"incidence" ($a_{ij} = 1$ si i "connaît" ou "est relié à" j pour $i \in \{1, \dots, n\}$; $a_{ij} = 0$ sinon) ; interprétation.
- Recherche de l'inverse d'une matrice 2×2 : position du problème, exemples simples.
- Puissances d'une matrice diagonale et exemple d'application.

4. Applications

Exemple 1 : portrait scolaire d'élèves

On pourra dresser le "portrait" scolaire d'un élève, noté en cinq matières, sous la forme d'un vecteur à 5 éléments représentant la moyenne en fin de premier trimestre dans les cinq matières ; deux autres vecteurs représentent cet élève au deuxième et troisième trimestre et la moyenne de ces trois vecteurs constituera le portrait scolaire annuel de cet élève.

Cet exemple pourra être repris pour avoir le portrait d'un groupe d'élèves. Ainsi, pour un groupe de 15 élèves, on pourra le représenter d'abord par le tableau 15×6 ci-dessous, où l'on a légendé les lignes et les colonnes, puis simplement par la matrice M .

	M	A	Φ	F	EPS
E1	16	16	6	14	18
E2	12	15	5	9	14
E3	11	16	2	6	14
E4	18	13	2	10	15
E5	5	13	6	5	10
E6	13	16	6	10	15
E7	15	14	6	11	14
E8	16	15	6	11	15
E9	20	14	6	13	16
E10	12	15	3	8	14
E11	9	18	5	7	15
E12	12	17	2	7	15
E13	12	15	3	7	14
E14	8	16	6	7	13
E15	15	17	6	15	19

$$M = \begin{pmatrix} 16 & 16 & 6 & 14 & 18 \\ 12 & 15 & 5 & 9 & 14 \\ 11 & 16 & 2 & 6 & 14 \\ 18 & 13 & 2 & 10 & 15 \\ 5 & 13 & 6 & 5 & 10 \\ 13 & 16 & 6 & 10 & 15 \\ 15 & 14 & 6 & 11 & 14 \\ 16 & 15 & 6 & 11 & 15 \\ 20 & 14 & 6 & 13 & 16 \\ 12 & 15 & 3 & 8 & 14 \\ 9 & 18 & 5 & 7 & 15 \\ 12 & 17 & 2 & 7 & 15 \\ 12 & 15 & 3 & 7 & 14 \\ 8 & 16 & 6 & 7 & 13 \\ 15 & 17 & 6 & 15 & 19 \end{pmatrix}$$

Si les matières envisagées ont des coefficients respectifs 4, 3, 2, 1, 1, on peut définir la matrice-colonne P des coefficients, puis la matrice-colonne P' des coefficients pondérés :

$$P = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \\ 1 \end{pmatrix} \text{ et } P' = \frac{1}{11} P$$

La matrice produit $R = M \times P'$ donne alors les notes finales des quinze élèves.

(Remarque : si M est la matrice-ligne correspondant à un seul élève, $M \times P'$ donne la note moyenne de cet élève.)

$$R = \begin{pmatrix} 14,182 \\ 11,574 \\ 10,465 \\ 12,682 \\ 7,810 \\ 12,570 \\ 12,835 \\ 13,539 \\ 14,784 \\ 10,997 \\ 10,894 \\ 11,598 \\ 10,863 \\ 10,033 \\ 14,273 \end{pmatrix}$$

Exemple 2 : Portrait électoral

On dispose, pour une ville de 2800 habitants, pour 2 tranches d'âge (J : moins de 45 ans ; NJ : plus de 46 ans) et trois modes de vie (H1, H2, H3), des pourcentages de vote pour trois candidats A,B,C au poste de maire ;

$$H1 = \begin{pmatrix} 30 & 50 & 20 \\ 25 & 40 & 35 \end{pmatrix} \quad H2 = \begin{pmatrix} 25 & 55 & 20 \\ 20 & 50 & 35 \end{pmatrix} \quad H3 = \begin{pmatrix} 50 & 30 & 20 \\ 40 & 40 & 20 \end{pmatrix}$$

Soient K_i les matrices définies par : $K_i = H_i/100$, $i=1,2,3$.

Les répartitions entre jeunes et non jeunes selon les modes de vie H1, H2, H3 sont respectivement :

$$\begin{pmatrix} 100 & 100 \end{pmatrix} \quad \begin{pmatrix} 1000 & 500 \end{pmatrix} \quad \begin{pmatrix} 100 & 1000 \end{pmatrix}$$

Considérons les trois matrices suivantes :

$$M1 = \begin{pmatrix} 100 & 0 \\ 0 & 100 \end{pmatrix} \quad M2 = \begin{pmatrix} 1000 & 0 \\ 0 & 500 \end{pmatrix} \quad M3 = \begin{pmatrix} 100 & 0 \\ 0 & 1000 \end{pmatrix}$$

Interpréter les termes des matrices ($M_i \times K_i$), pour $i=1,2,3$ ainsi que ceux de la somme S de ces trois matrices et ceux de la matrice $D \times S$, où :

$$D = \begin{pmatrix} 1/1200 & 0 \\ 0 & 1/1600 \end{pmatrix}$$

En étudiant cette élection, l'équipe du candidat A dit que les jeunes ont mieux voté pour A que les non jeunes et, au vu des résultats globaux, les candidats B et C réfutent ce commentaire puisque parmi les 1200 jeunes, 27,5% votent pour A alors que parmi les 1600 non jeunes, 32,8% votent pour A. Qu'en pensez-vous ?

Exemple 3 : Evolution d'un système

Dans une île, les mouvements de population peuvent être modélisés ainsi : chaque année, 40% des habitants de la capitale quittent celle-ci tandis que 20% des habitants du reste de l'île viennent y habiter ; on néglige les autres échanges. Partant d'une population (c_0, b_0) , quelle est la population deux ans après ? On introduira l'écriture matricielle :

$$\begin{pmatrix} c_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} 0,6 & 0,2 \\ 0,4 & 0,8 \end{pmatrix} \begin{pmatrix} c_0 \\ b_0 \end{pmatrix} \text{ soit } E_1 = A \times E_0.$$

E_i représentant l'état de la population la i -ème année, on montre que $E_2 = A \times A \times E_0 = A^2 \times E_0$, puis que $E_3 = A^3 \times E_0$, etc., en faisant un produit de matrices carrées ligne par colonne pour calculer les puissances successives d'une matrice ; on pourra étudier l'évolution sur n années ($n < 30$) suivant certains états initiaux et chercher pour quels états initiaux l'état du système est stable, en réfléchissant sur cette notion de stabilité (le nombre d'habitants respectifs de la capitale et du reste de l'île sont stables au cours du temps mais il y a toujours des gens qui partent de la capitale tandis que d'autres arrivent). On montrera que si un état d'équilibre s'installe, c'est un état stable. Cette notion de stabilité n'est pas au programme de 1^{ère} : il s'agit d'une première prise de contact ; elle a un sens intuitif très fort dont il serait dommage de se priver, et qui motive l'étude proposée.

Exemple 4 : Les boîtes de céréales

Trois sortes d'images sont réparties en proportions égales dans des boîtes de céréales.

Un inspecteur des fraudes, ayant observé ce qui se passait pour 1000 personnes achetant chaque semaine une boîte de céréale et voyant qu'au bout de 12 semaines, 15 personnes n'avaient que deux des trois images, a déclaré mensongère la publicité « les images sont également réparties dans les paquets ». Que penser de ce qu'il dit ?

Considérons donc 1000 personnes qui, chaque semaine, achètent exactement une boîte. Au bout d'une semaine, elles auront toutes une image. L'état E_n du système la $n^{\text{ième}}$ semaine est un vecteur colonne à trois lignes : la première composante est le nombre a_n de personnes ayant une seule sorte d'images, la seconde est le nombre b_n de personnes ayant exactement deux images distinctes, la troisième est le nombre c_n de personnes ayant les trois images. On a : $a_1 = 1000$, $b_1 = 0$, $c_1 = 0$. On admettra, à partir d'une justification heuristique, que pour $n > 1$:

$$\begin{aligned} a_n &= (1/3) a_{n-1} \\ b_n &= (2/3) a_{n-1} + (2/3) b_{n-1} \\ c_n &= (1/3) b_{n-1} + c_{n-1} \end{aligned}$$

Écrire ces trois équations à l'aide d'une seule équation matricielle et observer l'évolution du système sur 12 semaines. On utilisera la calculatrice pour faire les calculs.

Commenter l'accusation de l'inspecteur des fraudes.

Ce type de problème pourra être repris dans l'option de terminale.

Annexe

Note pour les professeurs, à propos des diagrammes en boîtes.

1-Quantiles et diagrammes en boîtes

En statistique, pour toute série numérique de données à valeurs dans un intervalle I , on définit la fonction quantile Q , de $[0,1]$ dans I , par :

$$Q(u) = \inf\{x, F(x) \geq u\},$$

où $F(x)$ désigne la fréquence des éléments de la série inférieurs ou égaux à x .

Soient $a_1, \dots, a_r$ les valeurs prises par une série de taille n , ordonnées par ordre croissant ; la fonction F est discontinue et constante sur les intervalles $[a_i, a_{i+1}[$; sa représentation graphique est composée de segments horizontaux.

En pratique, en consultant la liste des nombres $\{F(a_1), \dots, F(a_r)\}$, il est aisé de déterminer un quantile. Cependant, pour programmer le calcul de Q , on utilise la propriété suivante :

Soit n la taille de la série ; si on ordonne la série par ordre croissant, $Q(u)$ est la valeur du terme de cette série dont l'indice est le plus petit entier supérieur ou égal à nu .

Dans le cadre de cette définition, les 3 quartiles sont $Q(0,25)$, $Q(0,50)$, $Q(0,75)$. Les 9 déciles sont les valeurs de $Q(i/10)$, $i=1 \dots 9$, les 99 centiles sont les valeurs de $Q(i/100)$, $i=1 \dots 99$. On définit assez souvent la médiane m par $m = Q(0,5)$: la médiane est alors le second quartile, le cinquième décile, le cinquantième centile, etc...

Mais de nombreux statisticiens, de nombreux logiciels (de qualité) et de nombreux media utilisent la définition suivante de la médiane d'une série :

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.

C'est la définition adoptée dans le programme de seconde. Les deux définitions, $Q(0,5)$ et celle-ci donnent en pratique, pour des séries à valeurs continues de grande taille, des résultats le plus souvent très proches.

La procédure qui consiste à tracer une courbe dite de fréquences cumulées croissante, continue, obtenue par interpolation linéaire à partir des valeurs $F(a_i)$ définies ci-dessus et à définir la médiane comme l'intersection de cette courbe avec la droite d'équation $y=0,5$, où avec une courbe analogue dite des fréquences cumulées décroissantes n'est pas une pratique usuelle en statistique et ne sera pas proposée au lycée.

Si des données sont regroupées en classe, on parle de classe médiane.

Dans l'enseignement secondaire :

Pour les quartiles, la définition proposée sera liée à la fonction quantile :

Premier quartile : c'est le plus petit élément q des valeurs des termes de la série tel qu'au moins 25% des données soient inférieures ou égales à q .

Troisième quartile : c'est le plus petit élément q' des valeurs des termes de la série, tel qu'au moins 75% des données soient inférieures ou égales à q' .

Certains logiciels prennent pour le premier quartile une définition analogue à la médiane : par exemple si $n=4r$, le premier quartile est la demi-somme des valeurs prises par le terme de rang r et le terme de rang $r+1$. Nous n'adopterons pas cette définition un peu marginale.

On ne définira que le premier le neuvième décile :

Premier décile : c'est le plus petit élément d des valeurs des termes de la série tel qu'au moins 10% des données soient inférieures ou égales à d

Neuvième décile : c'est le plus petit élément d' des valeurs des termes de la série, tel qu'au moins 90% des données soient inférieures ou égales à d' .

On pourra introduire les termes suivants :

Intervalle interquartile : intervalle dont les extrémités sont le premier et le troisième quartile

Intervalle interdécile : intervalle dont les extrémités sont le premier et le neuvième quartile

Ecart interquartile : longueur de l'intervalle interquartile, i.e. différence entre le troisième et le premier quartile.

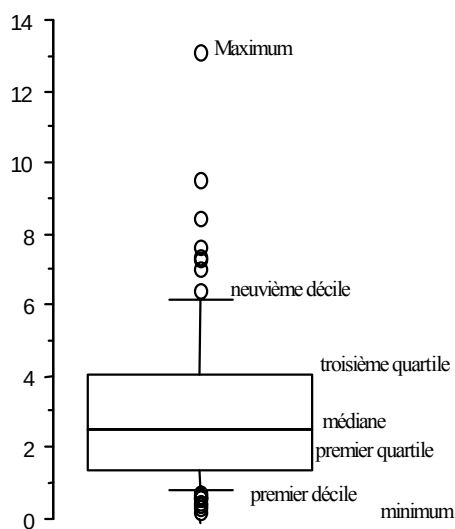
Ecart interdécile : longueur de l'intervalle interdécile, i.e. différence entre le neuvième et le premier décile.

Un abus de langage assez fréquent fait qu'on parle aussi d'intervalle interquartile au lieu d'écart interquartile : on évitera cet abus de langage au lycée.

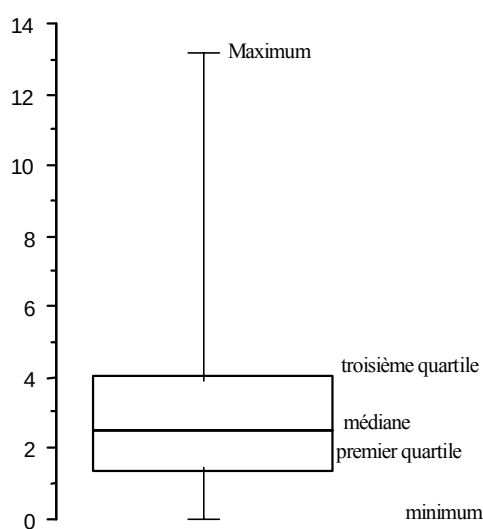
Remarque : Si les termes de la série subissent une transformation affine, $x \rightarrow ax+b$, où a est positif (ce qui est le cas lors d'un changement d'unité), les quantiles subissent la même transformation.

2-Diagrammes en boîtes

Ces diagrammes sont aussi appelés diagrammes de Tuckey, diagrammes à pattes ou à moustaches (whiskers plot). Il n'y a pas que le nom qui varie d'un logiciel à l'autre. Les deux situations les plus classiques sont représentées ci-dessous :



(1)



(2)

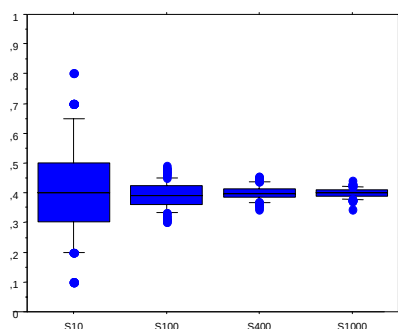
Nous choisirons de choisir « par défaut » la définition représentée graphiquement en (1), où figurent les premiers et neuvièmes déciles. Si une ou plusieurs valeurs extrêmes sortent résolument des limites du dessin, on indique dessous leurs valeurs sans les représenter. D'autres représentations peuvent être considérées, en précisant alors comment sont définis les « moustaches » de la boîte.

Néanmoins, les enseignants pourront utiliser des boîtes dont les extrémités sont les 1^{er} et 99^{ème} centiles, les valeurs extrêmes, etc. L'essentiel est d'avoir compris le principe : un jour d'examen, on demandera simplement à l'élève de spécifier en légende les éléments représentés.

Les premiers diagrammes en boîtes sont les diagrammes de Tuckey où la longueur des « moustaches » est 1,5 fois l'écart inter-quartile ; les diagrammes de Tuckey étaient utilisés dans des secteurs où les données peuvent le plus souvent être modélisées en utilisant une loi de Gauss ; dans ce cas, au niveau théorique, les extrémités des « moustaches » sont voisines du premier et 99ème centile : ces diagrammes étaient surtout utilisés pour détecter la présence de données exceptionnelles. On utilise aujourd'hui les diagrammes en boîtes pour représenter des distributions empiriques de données quelconques, non nécessairement symétriques autour de la moyenne, et le choix de moustaches de longueurs 1,5 fois l'écart interquartile ne se justifie plus.

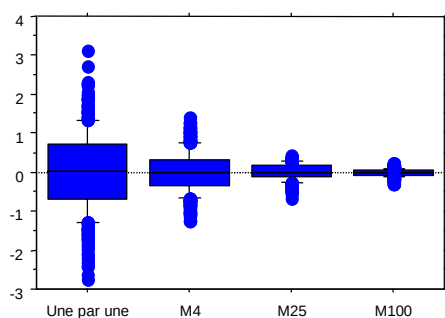
Les diagrammes en boîtes, comme les histogrammes, résument graphiquement une série ; l'idée de base est la suivante : au lieu de partager l'ensemble des valeurs possibles en segments égaux, on les partage en segments (quartile, déciles, centiles) qui contiennent une proportion prédéterminée des valeurs de la série. Les diagrammes en boîtes permettent de visualiser certains phénomènes et notamment de comparer plusieurs répartitions de valeurs. Ainsi, dans la figure ci-dessous, on a représenté les diagrammes en boîtes de :

- 100 simulations d'un sondage de taille 10 dans une population dont les individus sont codés 0 ou 1, la proportion de 1 étant ce qu'on cherche à déterminer (un sondage de taille n est ici le tirage au hasard –et avec remise– de n individus dans une population de taille N),
- 100 simulations d'un sondage de taille 100 dans la même population ,
- 100 simulations d'un sondage de taille 400 dans la même population ,
- 100 simulations d'un sondage de taille 1000 dans la même population .



La deuxième figure représente des mesures de hauteur d'eau dans un barrage par rapport à un niveau fixé : on met 100 appareils qui mesurent cette hauteur d'eau ; on a essayé quatre sortes d'appareils de mesure :

- ceux qui font une seule mesure,
- ceux qui font 4 mesures et donnent leur moyenne,
- ceux qui font 25 mesures et donnent leur moyenne,
- ceux qui font 100 mesures et donnent leur moyenne.

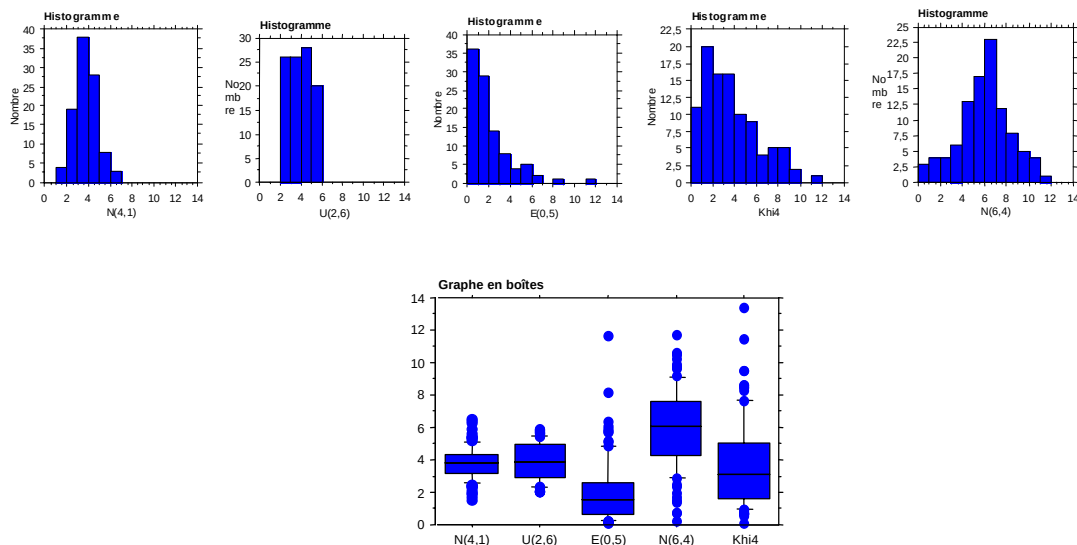


Les deux exemples situés ci-dessus sont spectaculaires et aisés à interpréter. Pour des séries de données quelconques, interpréter un diagramme en boîte demande un peu d'expérience et d'honnêteté pour ne pas transformer en affirmation théorique une observation lue sur un diagramme, que ce soit un histogramme ou un diagramme en boîte. Ci-dessous, nous présentons pour des séries de taille 100 simulées à partir de modèles classiquement utilisés divers résumés numériques et graphiques qu'on pourra s'exercer à lire :

- deux séries simulées à partir de lois de Gauss de moyenne 4 et 6 et de variance 1 et 4 et une série simulée à partir de la loi uniforme sur 1,6. Ces lois sont symétriques autour de leur moyenne : l'espérance et la médiane théorique coïncident et les graphiques théoriques sont symétriques.
- une série simulée à partir de la loi exponentielle d'espérance 2 et une série simulée à partir d'une loi du khi-deux à 4 degrés de libertés : ces lois de probabilité n'admettent pas de symétrie.

Statistiques descriptives

	Moy.	Dév. Std	Nombre	Minimum	Maximum	Médiane	Interquartile	10% Moy. élaguée
N(4,1)	3,782	,999	100	1,556	6,597	3,799	1,186	3,741
U(2,6)	3,927	1,153	100	2,006	5,927	3,881	1,999	3,932
E(0,5)	2,029	1,945	100	,055	11,633	1,562	1,926	1,706
N(6,4)	5,995	2,366	100	,185	11,762	6,056	3,376	6,021
Khi4	3,745	2,630	100	,095	13,430	3,085	3,428	3,423



Enfin, à ce propos et pour sa propre formation, l'enseignant pourra utiliser le logiciel SEL (voir §3). Plus précisément, il pourra :

- dans l'applet située à la page « diagrammes en boîte » du lexique voir comment fluctuent ces diagrammes lorsqu'on tire des échantillons au hasard dans une série de données réelles (tailles d'enfants de 6 ans).
- dans l'applet de simulations « diagrammes en bâtons, histogrammes et quantiles » superposer les histogrammes, fonctions de répartition, fonctions quantiles et diagrammes en boîtes de différentes lois classiques avec celles d'échantillons simulés.
- dans l'applet ajustement par quantiles, visualiser une technique classique d'ajustement de lois à des données.

3- Les logiciels SEL et SMEL

Ces logiciels ont pour origine la question suivante :

Comment se servir des outils proposés par le WEB pour l'auto-formation ou l'enseignement ?

Une réponse possible réside dans la mise à disposition de logiciels interactifs et dont la structure permet à la fois une mise à jour aisée et la transposition dans d'autres domaines.

Le domaine choisi est la statistique ; un premier logiciel SEL (statistique en ligne) est destiné à l'auto-formation des enseignants de sciences des lycées et collèges et concerne la statistique descriptive :

<http://www.inrialpes.fr/sel>

Un second logiciel, SMEL, qui englobe le premier, contient aussi des éléments de base de statistique inférentielle :

Ces deux logiciels sont structurés en quatre couches : une couche *articles* (26 articles dans SMEL), une couche *lexique*, une couche *cours*, une couche *simulation*. Dans les articles, les mots du lexique (environ 200 dans SMEL : loi normale, variance, loi de probabilité, etc.) sont cliquables : apparaît alors la page correspondante du lexique (page html), où le mot est brièvement défini et où figure dans certains cas une applet interactive (applet en java, le plus souvent construite sur un jeu de données réelles accessibles à partir de la page d'accueil du logiciel ; les graphiques comportent des curseurs mobiles et des fenêtres à contenus modifiables) ; dans cette page il y a de plus des fenêtres à menu déroulant :

- une fenêtre *Voir aussi* permet d'accéder directement à des mots en lien avec celui de la page ouverte,
- une fenêtre *Plus de détails* renvoie à un paragraphe du cours (où les mots du lexique sont aussi cliquables),
- une fenêtre *Lecture* renvoie aux articles ayant trait au contenu de la page.

Enfin, un renvoi à la couche simulations propose des applets de simulations (comparaison d'estimateurs, loi des grands nombres, analyse de variance, tests S, diagrammes, etc.).

Les pages de logiciel sont écrites automatiquement à partir d'un programme contenant tous les éléments à mettre dans cette page, sauf l'applet qui est spécifique du mot concerné dans la plupart des cas.

Chaque utilisateur se définira peu à peu des parcours de formation à l'intérieur du logiciel et grâce à l'interactivité pourra expérimenter soit à partir d'un jeu de données soit avec des données simulées. Ainsi, l'article « courbes de croissance » ouvre le chemin à un travail sur la loi de Gauss. Le néophyte pourra se familiariser avec les éléments de base de la statistique descriptive (quantiles, écart-type, divers diagrammes, etc.), voir ce que signifie la loi des grands nombres (et le contre-exemple de la loi de Student de paramètre 1), avoir une idée de ce qu'est l'analyse de la variance, et le moins néophyte pourra s'étonner devant les résultats proposés dans l'applet « variance biaisée et non biaisée »

Les logiciels SEL et SMEL sont téléchargeables et...gratuits !