HORS du PRET
4542
CRDP
E.M.
hors pret
I. R. E. M.
DE GRENOBLE
BIBLIOTHÈQUE

LANGAGE DES CHIFFRES

Sylviane GASQUET

LYCEE CHAMPOLLION – GRENOBLE

INTRODUCTION

Les fascicules "LANGAGE NUMERIQUE" et "LANGAGE GRAPHIQUE" sont complémentaires,mais cependant indépendants.

Ces deux brochures résument ou complètent le contenu des stages Mathématiques et Géographie, stages proposés depuis 3 ans au PAF de l'académie de Grenoble. Ils s'adressent à tous les consommateurs de chiffres et de graphiques.

Certaines parties sont des réponses à des demandes exprimées en stage, mais qui ne pouvaient être traitées face à un public varié (Géographes, Historiens, Matheux) sans risquer d'ennuyer certains (Pourcentages, maniement élémentaire d'une calculette, graphiques logarithmiques). A l'inverse, des thèmes un peu pointus n'ont pas été repris ici (Données corrigées des variations saisonnières, Elections à la proportionnelle, manipulation d'une matrice de données) pour ne pas alourdir.

Mais il est important aussi de dire que ce travail est écrit par un enseignant du secondaire, c'est à dire que les élèves sont continuellement présents, beaucoup de chapitres sont issus de séquences de classes (seconde, première A, terminales A,B,D). Un lecteur attentif pourra d'ailleurs percevoir mon cheminement : d'abord curiosité un peu naïve des manuels des autres disciplines, puis prise de conscience de l'apport possible pour "le cours de math", enfin maîtrise plus grande de certaines séquences abordées après plusieurs années de bain numérique! (l'analyse factorielle, par ex.)

Il ne faut pas gommer ce temps nécéssaire d'imprégnation avant de pouvoir utiliser réellement un domaine nouveau dans sa classe. Ces deux fascicules peuvent seulement permettre à un lecteur intéressé d'être un peu moins lent que moi ... Si certains d'entre vous ont déja lu des ouvrages intéressants et d'agréable lecture sur ce sujet, ce serait bien de faire parvenir les références ainsi qu'un bref commentaire, au CRDP : une édition ultérieure pourrait être complètée par une bibliographie plus riche. Le manque de temps est le principal ennemi de l'enseignant du secondaire lorsqu'il est d'un naturel curieux, peut-être pourrions nous nous aider à l'économiser ensemble?

Il me reste aussi à remercier tous ceux qui m'ont aidé dans ce travail:

- H.Chamussy de l'Institut de Géographie Alpine

- Des collègues du lycée Champollion, économistes, géographes ou philosophes,

- les participants aux stages de Valence, Annecy et Grenoble

- et mes élèves, particulièrement les élèves de seconde, dont la curiosité est réjouissante, même (surtout) lorsqu'ils arrivent affublés de l'étiquette "faible en math",

- enfin, je remercie tout spécialement M.Ruffieux, chargé d'Etudes au Rectorat de Grenoble, pour m'avoir permis de m'associer à ses travaux et m'avoir fait découvrir la manipulation numérique en vraie grandeur. Notre collaboration se poursuit, il serait dommage d'interrompre notre match amical aux erreurs! (Ne m'a-t-il pas surprise en flagrant délit d'addition illégitime de pourcentages!)

Le 3 janvier 87,

Sylviane Gasquet,

Lycée Champollion. Grenoble.
IREM (en heures supplémentaires...)

FASCICULE LANGAGE DES CHIFFRES

SOMMAIRE

*** I . Apprivoiser les chiffres. . . . . page 5

Le langage numérique n'est pas inné. Il faut découvrir ses pièges et ses paradoxes.

*** II . Pourcentages : Attention, prudence ! page 21

Il faudrait simplifier et coordonner nos façons de faire calculer des pourcentages. Peut-on penser les pourcentages comme des outils de communication ? Pour finir, un mode d'emploi sur calculette.

*** III . Cinquante millions de consommateurs de Statistiques:page 47

Découvrir l'état d'esprit des statistiques et non les formules.

*** IV . Initiation à l'analyse factorielle. . . . .page 62

Tableaux numériques et tableau noir : apprendre à lire! Où l'on utilise le produit scalaire...sans le savoir!

*** V . Faire avouer les chiffres. . . . . page 76

Comment un graphique peut faire dire à des chiffres ce qu'ils gardaient secret...

*** VI . Démystifier le chiffre . . . . . page 84

Oui ou non, existe-t-il un double langage des chiffres, ou bien tout simplement un langage qu'il nous faut découvrir? Pourquoi cet attrait du chiffre ?

I — APPRIVOISER LES CHIFFRES

*** Le siècle des "chiffres".

Si, au cinquième siècle, l'empire romain fut envahi par les HUNS, on peut dire que le vingtième siècle aura vu une invasion d'une violence plus pernicieuse, celle des pays industrialisés par les CHIFFRES.

Qu'ils soient associés à un jeu paisible

Le yoyo des chiffres

LE MONDE — Mardi 4 juin 1985 •••

ou plus clairement présentés comme l'arme nouvelle:

Bataille de chiffres sur l'immigration

Entre la gauche et la droite,
le débat s'est engagé de la pire manière,
à coups de slogans
et de statistiques contestées

— Mercredi 13 novembre 1985 •••

une chose est sûre, ILS sont là! Que la bataille de chiffres remplace celle des chars, soyons-en conscient :l'assaut est donné.

Précisons d'emblée que nous optons résolument pour le vocabulaire du français. Si en mathématique le mot chiffre est strictement réservé aux dix signes de base de la numération, et si 1986, assemblage de quatre chiffres s'appelle un nombre, il n'en est pas de même dans le langage courant, sans doute parce que verbe et adjectif sont dérivés du mot chiffre:

- Donnez-moi des informations chiffrées....

- Chiffrez-moi le montant probable de vos dépenses...

Cependant, on dira .plutôt que "les grands nombres, ou les très petits nombres dépassent notre entendement...", comme si ,aux extrêmes (qui n'en sont pas!) les chiffres devenaient nombres...

Voici, pour illustrer notre propos, un article distribué à des élèves de première A:

Le Monde. 2 Mai 85.

La France des femmes seules

La solitude gagne dans la société française. C'est une des évolutions les plus marquantes qui émergent des résultats du recensement de 1982 analysés dans le dernier numéro d'*Economie et statistique*, la revue de l'INSEE (1). Les plus touchées sont les femmes, vouées en nombre croissant à passer seules des périodes de plus en plus longues, surtout à la fin de leur vie.

Entre les deux recensements de 1975 et 1982, le nombre des personnes vivant seules a augmenté de moitié, passant de 3,2 millions environ à plus de 4,8 millions, soit 9 % de la population des ménages, alors que, de 1946 à 1975, la proportion n'avait guère dépassé 6 %.

Plus confortablement

Pour 62 %,ces solitaires ont plus de cinquante-cinq ans. Les personnes âgées habitent de plus en plus fréquemment seules : c'était le cas de 36,5 % des personnes de soixante-quinze ans et plus en 1982, contre 28,5 % vingt ans plus tôt. Elles étaient devenues deux fois plus nombreuses que celles qui vivaient chez des proches, notamment chez leurs enfants (19,3 %, contre 34,3 % en 1962).

Evolution contraire à celle des jeunes : un nombre croissant de dix-huit-vingt-quatre ans restent chez leurs parents parce qu'ils font des études de plus en plus longues, et surtout parce qu'ils ont de plus en plus de mal à trouver un emploi.

Cet isolement des personnes âgées, accompagné et facilité par la revalorisation des retraites, ne doit pas être décrit sous des couleurs uniformément sombres. En effet, les « solitaires âgés » vivent plus confortablement qu'autrefois : en 1982, 65 % des personnes seules de soixante-cinq ans et plus ont le téléphone, contre 19 % en 1975; 41 % disposent à la fois d'installations sanitaires, de WC intérieurs et du chauffage central, contre 24 % sept ans plus tôt.

Les quatre cinquièmes de ces « solitaires âgés » sont des femmes essentiellement par suite de la *« surmortalité »* masculine : on comptait ainsi, en 1982, cinq fois plus de veuves (2 700 000) que de veufs (540 000). Après soixante ans, celles-ci sont plus nombreuses que les femmes mariées, alors que, au même âge, les hommes mariés restent six fois plus nombreux que les veufs. Il faut encore y ajouter les célibataires et les divorcées. Résultat : à partir de soixante-treize ans, la solitude est le mode de vie le plus répandu chez les femmes : plus que le mariage, que la vie chez des proches et — jusqu'à quatre-vingt-dix ans tout au moins — que la résidence dans une maison de retraite ou un hospice.

La fréquence varie cependant selon les régions : la proportion de solitaires chez les femmes âgées de quatre-vingts à quatre-vingt-quatre ans dépasse 60 % dans l'Ile-de-France, la Champagne, la Bourgogne, le Nord, la Normandie, le Centre, les régions Rhône-Alpes et Provence-Côte d'Azur. En revanche, elle est inférieure à 50 % dans tout le Sud-Ouest, en Auvergne et dans le Limousin.

Mais l'isolement devient plus fréquent aussi chez les femmes d'âge actif (vingt-soixante ans), avec l'extension du célibat et la multiplication des divorces. La proportion de célibataires « définitives » chez les femmes de quarante-cinq ans a été multipliée par sept de 1962 à 1982. Elle approche 20 % pour les femmes ayant fait des études supérieures. Tout se passe comme si les femmes diplômées et actives avaient plus de mal à se marier que les autres, et étaient plus que les autres, et plus que les hommes, vouées à vivre seules malgré le développement de l'union libre. Ainsi, avant trente-quatre ans, 22,1 % des femmes cadres supérieurs et 16,1 % de celles exercant une profession « intermédiaire » (cadres moyens) vivent seules, contre respectivement 15,2 % et 11,2 % des hommes.

Quant au divorce, il tend à élargir une autre catégorie de solitaires, ou presque : les femmes seules avec un ou deux enfants, dont le nombre a augmenté de plus de 20 % entre 1972 et 1982. Parmi elles, les divorcées sont près de 400 000 ; elles représentent à elles seules la moitié des familles « monoparentales » (au lieu d'un tiers en 1975) et les célibataires 14,5 %.

GUY HERZLICH.

(1) *Economie et statistique* n° 175, mars 1985.

" Voilà, en effet, que la table est mise pour le condéssirable , surpressionnant, planturesque festin.... (Les chiffres) dé-bordent à présent, se resplandent, claboussent et pullugrouillent. On les voit dans certains documents, promener leur monumental foisonnement à travers les luxurieuses circonstances, comme une explosion prodigieuse venant ordonner l'anarchie."

Je n'ai pas résisté : le texte précédent est emprunté à A. DELEDICQ, (préface de son livre "Mathématiques buissonnières") mais en osant une substitution: il parlait des mathématiques et non des chiffres.

Bien sûr, telle n'a pas été la réaction des élèves de première. Ce fut presque unanimement : on n'y comprend rien, on n'est pas assez fort (en math) pour cela... Or un élève qui, utilisant les tableaux numériques donnés dans les sujets de bac, ferait un devoir de ce style aurait immanquablement une note lamentable, car de ce genre d'article on ne retient rien...sauf ce que l'on savait déjà! A titre d'exercice, j'ai donné à compter le nombre de référentiels différents utilisés....(C'est-à-dire les ensembles cités. Exemple : les femmes de 80 à 84 ans; les femmes de plus de 73 ans, etc...)

Dans la suite de cette première partie consacrée à l'information chiffrée, nous allons donner des exemples de travaux faits en classe (seconde, 1ère A ou B, Terminale B).

*** Initiation à l'information numérique.

Voici un exemple proposé en seconde, sous forme de problème ouvert. On suppose qu'un journaliste local doit "rédiger un papier" à propos des données suivantes:

Lycées	Nombre de candidats	Nombre de reçus au Bac
A	100	65
B	150	90
C	250	180
D	250	150
E	150	110
F	200	130
G	350	210

La première réaction des élèves est de calculer les pourcentages et de les disposer dans une quatrième colonne.

- Leur manière de calculer lesdits pourcentages est trés variable et parfois erronée,mais nous reviendrons sur ce sujet.

- pourquoi ces calculs?
- pour comparer, pour classer...

Il s'agit alors de faire prendre conscience aux élèves de ce qu'ils font en passant aux "pourcentages":

- A partir des "données brutes", ils fabriquent un indice, qui est un chiffre artificiel (issu du choix d'une opération).

- Il est vrai que chaque lycée était connu par deux nombres et qu'on ne peut classer aisément des "couples" (découverte d'un ensemble non ordonné "naturellement")

- en passant des données brutes (qui sont la matière première) à l'indice fabriqué, on perd de l'information (ici, la taille des établissements).

- Et que mesure exactement l'indice fabriqué? La controverse s'instaure avant que l'accord ne se fasse... Ce n'est pas la réussite du lycée A ou B, mais la réussite des terminales du lycée, ce qui n'est pas la même chose....

Il s'agit ici de faire sentir aux élèves l'importance de l'ensemble de référence. Un pourcentage ne devrait jamais être cité sans que le référentiel soit clairement nommé.

D'ailleurs, même dans ce cas, l'erreur n'est pas rare, en voici un exemple drôle:

Une nouvelle qui étonne les médecins aux entretiens de Bichat

Le nouveau mâle hystérique en représentation

Selon le professeur Henry Loo, de l'hôpital Sainte-Anne, un tiers des hommes est concerné par l'hystérie. Particulière et originale, celle-ci est difficile à diagnostiquer et ne doit rien à sa concurrente féminine.

Une hystérie sur trois concerne un homme : la nouvelle semble avoir passablement ébranlé le parterre de médecins venus, au 40e Entretiens de Bichat, recueillir la bonne parole médicale. Le professeur Henry Loo, de l'hôpital Sainte-Anne, a en effet annoncé que, selon les statistiques épidémiologiques des centres de sélection militaire.

Reste que l'ampleur du mal n'est toujours pas évaluée, faute de diagnostic précis. En 1981, le docteur Darcourt évaluait à 34 % la répartition masculine de ces personnalités hystéri-

LIBERATION • VENDREDI 3 OCTOBRE 1986 **29**

Dans le "chapeau" de l'article, l'ensemble de référence est l'ensemble des hommes, dont un tiers serait hystérique....ce qui aurait tout de même dû faire sursauter le journaliste!

Mais, au début de l'article, l'ensemble de référence est l'ensemble des malades hystériques (dont on ignore totalement l'effectif, ne serait-ce que pour se faire une petite idée...), et parmi ces malades, un sur trois est un homme....ce qui est heureusement tout à fait différent de ce qu'annonçait le sous-titre!

Mais revenons au thème du bilan d'un lycée, thème qui se place naturellement dans les préoccupations des élèves:

- pour parler de la réussite d'un lycée, quelles données faudrait-il collecter?

- Delphine propose un chiffre unique pour pouvoir conclure sur la région étudiée: la moyenne (arithmétique) des pourcentages. Le maître élude cette proposition, qui a séduit tout le groupe, en disant qu'on reprendrait l'idée plus tard! En effet, un rapide calcul à la machine lui montre que, vu les données, la moyenne arithmétique est peu différente de la moyenne pondérée...et donc qu'il serait vain, sur cet exemple, d'essayer de mettre en évidence la nécéssité de coefficients de pondération...

... et pour changer de cap, le maître demande si on ne peut pas trouver une illustration graphique des données brutes afin de conserver toute l'information...

*** Faire SENTIR la nécéssité d'une moyenne pondérée.

Voici le travail à la maison proposé à la classe de seconde suite au travail sur les données relatives à la réussite au bac.

....Sur une idée de Delphine : calculons la moyenne.

Dans un département imaginaire, il y a les lycées suivants:

Patronymes	Nombre de candidats	Nombre de reçus au bac	Taux de réussite en terminales.
Dupont	50	40	
Duparc	60	55	
Duval	75	62	
Dubois	75	65	
Dutour	80	70	
Dubuisson	400	180	

Remplir la dernière colonne. Puis, comme Delphine le proposait Mardi, faire la "moyenne des pourcentages".

Ayant fait ce calcul, l'inspecteur d'académie est trés satisfait du résultat, il le dit à FR4.

Mais, calcule aussi le nombre total de candidats dans le département, le nombre total de reçus, et finalement le taux de réussite du département. Ce résultat est-il le même que la "moyenne des taux"?

Rédige ce que tu penses à ce propos sous la forme d'une lettre à l'inspecteur d'académie.

. . . . . .

Bien sûr, les données ont été choisies ici pour que l'écart soit grand entre la moyenne arithmétique et le taux réel. La première vaut 78% et le taux 63%.

Voici quelques extraits des lettres reçues... Précisons qu'il s'agit, exceptionnellement cette année, d'une seconde expérimentale non différenciée : les élèves ont été regroupés sur la base de leur allergie aux mathématiques...(les deux tiers suivent une option de quatre heures de dessin ou de musique)

" ... Dans votre calcul, vous oubliez de mettre en valeur l'importance des établissements et d'en tenir compte. Ainsi le lycée Dubuisson,quatre fois plus important que celui Dupont et qui a un taux trés faible se trouve noyé dans les chiffres annoncés."

Martine,

"Le résultat de l'enquète sur le taux de réussite au bac en terminale était en effet très réjouissant.

Malheureusement certaines personnes animées par une curiosité toute légitime se sont amusées à retraiter les données brutes.(...)

Auquel de ces chiffres faut-il se fier ? En fait l'erreur se situe dans le premier résultat: en effet tous les lycées n'ayant pas le même nombre de candidats, les pourcentages ne sont pas proportionnels et on ne peut les additionner dans ce cas....."

Valérie,

" .... En effet on ne peut pas effectuer une moyenne générale vue la disproportion du nombre d'effectifs de chaque lycée notemment visible entre le lycée Dupont et Dubuisson. Ce "taux de réussite" du département n'est donc pas précis et en l'occurence faux."

Delphine,

".... Ce phénomène étrange (la différence des deux résultats) s'explique tout simplement par le fait que les meilleurs taux de réussite au bac sont obtenus par des petits lycées présentant entre 50 et 80 candidats, alors que le lycée Dubuisson, lycée au centre de notre coquette préfecture, Ambitions sur Isère, obtient lui, un nombre de reçus tout à fait moyen tout en présentant 5 à 8 fois plus de candidats au bac que les autres lycées du département."

Emmanuel,

D'autres élèves se sont bornés à constater la différence entre les deux résultats...et à la faire connaître à l'inspecteur,sans chercher d'explications.

(Et deux élèves...ont même trouvé pareil par les deux modes de calculs!)

Lors du débat en classe, Florence émet l'idée qu'il faut donc toujours connaître les données brutes. Il s'ensuit alors l'idée de mettre un coefficient proportionnel à la taille de l'établissement.C'est le coefficient d'importance du lycée.

Ainsi les données suivantes permettent de trouver le taux réel de réussite dans le département:

Patronymes	Coefficient d'importance	Taux de réussite en terminales.
Dupont	6,7 %	80 %
Duparc	8,1 %	91,7 %
Duval	10,1 %	82,7 %
Dubois	10,1 %	86,7 %
Dutour	10,8 %	87,5 %
Dubuisson	54,1 %	45 %

On retrouve la même notion que les coefficients donnés aux notes d'un examen: chaque lycée a sa propre note de réussite, mais il pèse plus ou moins lourd dans le département suivant sa taille.

D'où la moyenne pondérée:

$$\frac{80.6{,}7 + 91{,}7.8{,}1 + 82{,}7.10{,}1 + \ldots}{100}$$

100 étant la somme des coefficients, aux erreurs d'arrondis près.

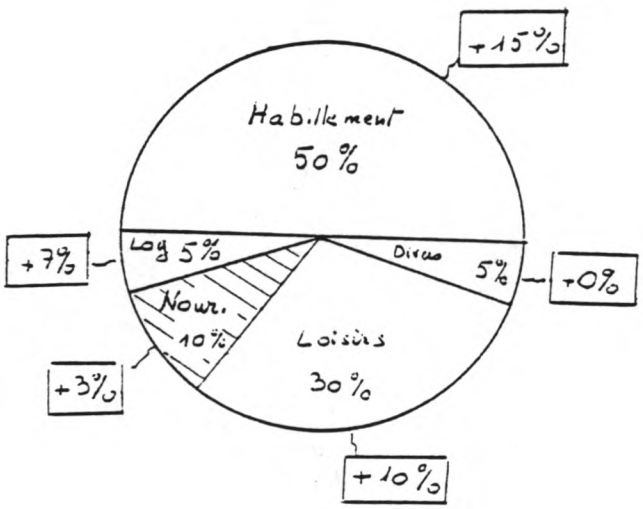
Les mots choisis ("peser plus ou moins lourd") ne sont pas anodins. Même dans cette classe "en difficulté", une élève sentira le lien avec le barycentre d'une plaque homogène (choix d'un coefficient proportionnel à l'aire, donc à l'importance de la portion de plaque). Je montre à Florence que la formule de la moyenne pondérée est la même que celle des coordonnées du barycentre. Le barycentre est une "moyenne pondérée de points".

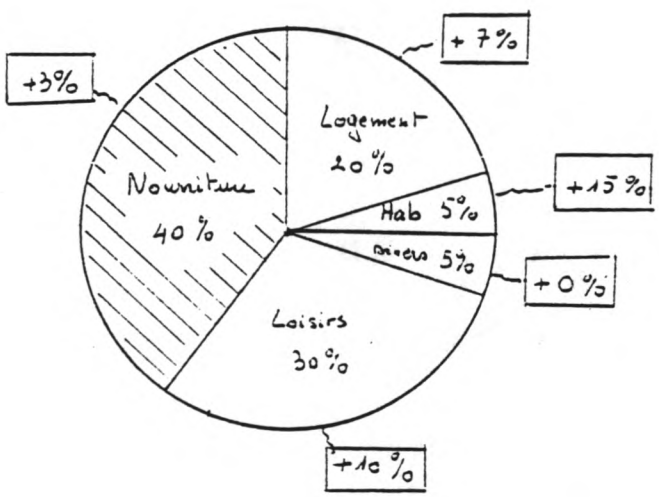
Pour la majorité de la classe, cependant, un autre modèle sera utile:

Voici la répartition des dépenses de deux personnes que les élèves se plaisent à imaginer.... En général, le "coquet" est vu au féminin, tandis que le "gourmand" est plutôt masculin!

Je tiens à signaler que ces budjets sont évidemment caricaturaux, mais ce procédé "pédagogique" me sembla très efficace sur moi-même lorsque je lus "Attention Statistiques!", et donc je me le suis approprié: il serait dommage que les méthodes astucieuses ne se répandent pas...

La surprise, pour les élèves, est de voir qu'on ne connaît ici aucune donnée brute, et il leur semble difficile de croire qu'on va pouvoir dire, au vu des augmentations du prix de la vie,de quel pourcentage augmentera le budget de chaque personne...





Le coquet...

Le Goulu!

Ces deux personnes se trouvent face aux mêmes augmentations de la vie:

- Habillement + 15%
- Logement + 7%
- Nourriture + 3%
- Loisirs + 10%
- Divers 0%

Mais les élèves sentent bien que, l'habillement augmentant de 15% et la nourriture de 3%, les répercussions seront plus dures pour le "coquet"...

Ils sentent alors que la moyenne (arithmétique) des augmentations ne représente rien pour ces deux personnes. La notion de calcul possible certes, mais illégitime, commence à faire son chemin.

Ils ne proposeront pas la moyenne pondérée, mais à l'inverse Karine saura dire que si une personne avait une répartition équilibrée (soit 20 % pour chaque rubrique) alors "ça serait bien la moyenne ordinaire des augmentations".

Karine perçoit d'ailleurs très bien que l'augmentation du budget d'une personne dépend, à parts égales, à la fois des augmentations extérieures, et à la fois de la répartition du budget. Elle pense pouvoir inventer le budget d'une personne plus coquette que le "goulu" mais dont cependant le budget augmenterait moins, en pourcentage,que celui du goulu...(et pourtant l'habillement augmente 5 fois plus!).

J'ai détaillé ce travail pour deux raisons.

Tout d'abord parce qu'il me semble invraisemblable que, dans notre société actuelle, l'école ne forme pas les jeunes au langage numérique alors qu'ils le rencontrent dès qu'ils ouvrent un journal, la télévision ou leur livre de géographie....

Ensuite, en dehors du contenu, pour bien montrer ce que j'appelle "faire sentir" la nécessité de pondérer une moyenne.

Peut-être (et même sûrement!) les élèves oublieront-ils comment on pondère... mais encore plus important me semble être de ne pas accepter n'importe quel mode de calculs... Nos élèves, comme nous-mêmes, ne seront pas des "fabriquants" d'informations numériques, ils en seront des consommateurs obligés. Il me suffirait, à l'issue de cette séquence, que les élèves puissent refuser la moyenne arithmétique dans les cas où elle n'est porteuse d'aucune signification.

Ce genre de réflexion peut suffire. En tout état de cause, il doit précéder l'arrivée d'une "formule", sur la moyenne pondérée dans le cas qui nous occupe ici.

*** Conjecturer à partir d'un chiffre...

Il est important de faire sentir qu'un seul chiffre ne peut résumer une réalité. Au contraire, il faut apprendre à "se poser des questions" (quelles sont les diverses causes possibles amenant ce résultat?) pour aller chercher d'autres chiffres permettant d'appuyer ou au contraire de réfuter une hypothèse possible.

- A - Dans ce lycée, le pourcentage de réussite au bac a augmenté. Que peut-on penser ?

D'abord bien énoncer le référentiel: c'est l'ensemble des élèves de terminale du lycée.

Le pourcentage est un rapport. On nous dit qu'il a augmenté.

Alors,il se peut

- que pour un nombre identique de candidats, il y a eu davantage de reçus.
- qu'il n'y ait pas plus de reçus que l'année précédente, mais beaucoup moins de candidats, par exemple parce qu'il y a eu plus de redoublements en première ?
- qu'il y ait encore moins de reçus que l'an passé... mais aussi beaucoup moins de candidats?

Toute évolution d'un pourcentage doit nous alerter de deux façons. S'il s'agit d'une augmentation, est-ce le numérateur seul qui a augmenté? est-ce le dénominateur seul qui a diminué? Les deux raisons sont peut-être intervenues en même temps?

Or l'usage du pourcentage, qui nous fait entendre à l'oreille un immuable "POUR CENT", incite à oublier que le dénominateur peut lui aussi varier...

Il faut donc aller s'enquérir d'un autre chiffre, celui de l'effectif des candidats pour les deux années. (ou éventuellement de savoir seulement qu'il est demeuré pratiquement le même).

- B - En France, depuis une dizaine d'années, le Revenu moyen des Agriculteurs a augmenté.

- les revenus de tous les agriculteurs ont augmenté ?

- certains agriculteurs ont vu augmenter leurs revenus et d'autres l'ont vu baisser,le total des augmentations étant supérieur aux baisses ?

- aucun revenu agricole n'a augmenté...les petits revenus ont disparu (abandon des petites exploitations), et donc la moyenne a monté.

Donc il serait bon de voir si le nombre d'exploitations est demeuré à peu près constant....

C'est en effet la dernière supposition qui s'avère être la réalité. Sans doute sera-t-on étonné, comme je le fus, d'apprendre qu'un élève de seconde proposa bien la troisième hypothèse. Mais Laurent reconnut qu'il avait surtout médité que, si je posais la question, il devait y avoir du paradoxe dans l'air! Il commença donc par se demander : est-il possible qu'aucun revenu n'augmente? dans quelles conditions cela est-il possible...

Il est bien entendu que lorsque l'évolution d'un tel chiffre est donné oralement, dans un débat télévisé, peu d'auditeurs pensent à ce troisième cas... Ce n'est déjà pas si mal s'ils ont le temps d'évoquer le second cas.

Nous reviendrons sans doute souvent sur l'idée qu'il est impossible de maîtriser l'information numérique lorsqu'elle est délivrée oralement. S'il y a trop de chiffres,on s'y perd, et s'il n'y en a qu'un on n'a pas le temps de conjecturer les diverses possibilités...Tout au plus peut-on se dire : méfiance!

- C - Passage d'une moyenne sur un ensemble...à la réalité dans des sous-ensembles...

Cet exemple, comme celui qui précède, est extrait du livre "ATTENTION, STATISTIQUES!" de Joseph Klatzmann, paru aux éditions La Découverte, Cahiers libres n° 405, livre dont la lecture est des plus toniques.

Une enquête menée dans la région du Nord nous apprend que le taux de mortalité infantile est plus élevé dans les ménages où la mère ne travaille pas.

On pourrait "absorber" ce résultat comme un information plate, sans mystère... Et pourtant ,l'auteur ajoute que ce résultat est inversé chez les cadres supérieurs, chez les cadres moyens, chez les ouvriers.... Sans doute aura-t-on du mal à croire la chose mathématiquement possible, mais dans TOUTES les catégories socio-professionnelles, le taux de mortalité infantile est au contraire plus faible dans les ménages où la mère ne travaille pas....

Voici l'explication , sous forme d'exemple simplifié et caricatural, proposé par J. Klatzmann, page 109. Nous allons pouvoir faire les calculs pour nous convaincre et aussi pour comprendre.

REPARTITION DE LA CONSOMMATION DE POMMES DE TERRE

Région 1: 10 agriculteurs consomment chacun 130 kg de pommes de terre par an; 90 non-agriculteurs en consomment 120 kg.

Région 2: 50 agriculteurs consomment chacun 90 kg de pommes de terre par an; 50 non-agriculteurs en consomment 80 kg.

Notons bien que dans chacune des deux régions, la consommation des agriculteurs est supérieure à celle des non agriculteurs.

Pourtant si on fait le calcul pour "l'ensemble du pays",les 60 agriculteurs consomment 5800kg, soit une moyenne d'environ 97 kg par personne. Et les 140 non-agriculteurs en consomment 14800,soit à peu près 106 kg par personne.

Les moyennes sont inversées! Ce sont les agriculteurs qui consomment le moins!

Il n'y a pas ici un chiffre faux et un chiffre vrai (comme pour la moyenne des pourcentages de réussite au bac), il y a les chiffres pour l'ensemble et les chiffres par région.

L'inversion de l'ordre vient du fait que dans la région 1, les "petits consommateurs" (pour leur région!) consomment en réalité plus que les "gros consommateurs" de la région 2. Et comme ils sont nombreux,ils font basculer l'ordre lorsqu'on fait le calcul sur l'ensemble des deux régions...

Il en est de même pour le taux de mortalité infantile. Une catégorie était sur-représentée dans la région du Nord, celle des familles de mineurs. Chez ces familles,le taux de mortalité infantile est aussi plus bas quand la mère ne travaille pas (et elles sont très nombreuses à ne pas travailler), mais ce taux plus bas relativement à leur propre CSP, est cependant bien plus élevé que les taux des ménages où la femme travaille pour les autres CSP, (à cause du manque d'hygiène et d'un plus faible niveau d'instruction).

Notons à l'intention des lecteurs géographes, que nous ignorons la date de cette enquête. Le livre de Klaztmann nous incite à réflèchir sur des résultats (en général assez anciens ou non datés), il n'est pas un livre de statistiques.

Retenons donc qu'un résultat vrai sur un ensemble ne l'est pas nécéssairement sur les diverses strates dans lesquelles on partage cet ensemble.

- D - Entre 1931 et 1936,la population active agricole française passe de 7,7 à 7,2 millions, soit environ 1oo ooo de moins par an.

Peut-on parler d'exode rural ?

Solution page 44 de Attention Statistiques!

(Cet exercice est utilisable en devoir surveillé : on attend de l'élève qu'il précise quelle donnée supplémentaire est indispensable pour pouvoir dire si, oui ou non ,il y a eu exode rural.)

*** Ne pas se laisser tenter par l'esprit d'addition...

Voici un exemple extrait du livre d'Albert Jacquard, " Moi et les autres" (collection Point,Virgule au Seuil), où l'on retrouve la méthode consistant à partir d'un exmple simple pour mieux saisir une réalité plus complexe:

ENCADRÉ Nº 3

LES PIÈGES DE L'ANALYSE DES VARIATIONS

En 1981, j'ai acheté chez mon libraire 20 livres de poche à 10 F et 20 livres d'art à 100 F, soit une dépense totale de 2 200 F.

En 1982, le prix des livres de poche a augmenté et atteint 15 F, en revanche celui des livres d'art est revenu à 90 F. J'ai acheté 60 livres de poche et 15 livres d'art, soit une dépense totale de 2 250 F.

Dans l'augmentation de ma dépense (+ 50 F), quelle est la part due à la variation des quantités, la part due à la variation des prix ?

Répondre semble facile :

— si les prix étaient restés ceux de 1981, ma dépense aurait été de $60 \times 10 + 15 \times 100 = 2\,100$: le changement des quantités a donc entraîné une diminution de 100 F ;

— si les quantités achetées avaient été les mêmes qu'en 1981, ma dépense aurait été de $20 \times 15 + 20 \times 90 = 2\,100$ F : le changement des prix a donc entraîné une diminution de 100 F.

Agissant séparément, la variation des prix aurait provoqué une diminution de ma dépense ; il en est de même pour la variation des quantités. Et pourtant j'ai dépensé 50 F de plus !

En fait, l'erreur vient de ce que les variations calculées en isolant chaque facteur n'ont absolument aucun sens. C'est la question posée « Quelle est la part due... » qui est absurde. Quelle que soit la technique utilisée pour y répondre, la réponse ne peut, elle aussi, qu'être absurde.

Comment certains psychologues croient-ils encore possible d'évoquer les « parts » de l'inné et de l'acquis dans l'intelligence ?

Il faut prendre soi-même son crayon pour constater ces bizarreries numériques...Les nombres ne se laissent pas apprivoiser si aisément que cela....

Le livre cité ici est en fait un livre de vulgarisation à l'usage des adolescents. Des enseignants curieux des jeux de chiffres et de génétique liront plutôt "Eloge de la différence", du même auteur. On peut y trouver des thèmes spécifiques pour la classe de terminale D, nécéssitant à la fois des connaissances de base de génétique et de calcul des probabilités.Nous préfèrons ici traiter d'exemples accessibles à toutes les classes de lycée.

Ayant souvent perçu une sorte de dédain des enseignants de mathématiques à l'égard des calculs numériques , je tiens ici à faire percevoir tout ce que ce genre de travail peut amener comme réflexion....On est loin des automatismes dérivée, tableau de variations etc. Ce qui est grave c'est que l'institution scolaire arrive à persuader les élèves eux-mêmes qu'une hiérarchie "immanente" existe entre les mathématiques "avec des lettres" et celles avec des chiffres. Ainsi une classe de première A finit-elle par avouer les causes de son rejet initial face au domaine numérique: les élèves pensaient que j'avais choisi ce thème parce que je les trouvais "trop bêtes" pour faire des mathématiques traditionnelles....

Nous verrons d'autres exemples, mais où le numérique rejoint le graphique.

Rappelons cependant que de nombreux exemples purement numériques sont détaillés dans un fascicule édité par le CRDP de Grenoble: Liaison Math - Economie.

Après ces exemples, nous allons dire quelques mots des pourcentages, en particulier de leur apprentissage.

II - POURCENTAGES : ATTENTION, PRUDENCE !

Nous distinguerons successivement deux types de pourcentage:

- celui qui rend compte du rapport entre une partie et un tout,c'est le pourcentage "synchronique",

- celui qui suit une évolution au cours du temps (augmentation, baisse), c'est l'aspect diachronique .

... I ... A PROPOS DU POURCENTAGE SYNCHRONIQUE.

*** 1.Ecriture d'un rapport en pourcentage.

- Tout rapport ne se traduit pas en pourcentage.

Ainsi une densité d'habitant ou une vitesse...

Règle 1 Un rapport se traduit en pourcentage si le numérateur dénombre les éléments d'un ensemble INCLUS dans l'ensemble dénombré par le dénominateur.

Pourcentage synchronique : $\frac{\text{sous-ensemble}}{\text{ensemble de référence}}$

Exemple: $\frac{\text{Nombre de reçus au bac dans le lycée M}}{\text{Nombre de candidats au bac dans ce même lycée.}}$

Ainsi le rapport $\frac{450}{600}$

a une écriture décimale 0,75 ET une écriture journalistique 75%

(laquelle rappelle d'ailleurs que le nombre décimal 0,75 correspond à la fraction décimale 75/100)

Règle 2 - Le pourcentage synchronique ne peut dépasser 1.

Autrement dit, on ne peut pas compter plusieurs fois un même élément lorsqu'on dénombre l'effectif de la partie concernée par le numérateur.

Cette règle est déja moins évidente à respecter, surtout lorsqu'on se lance en apprenti sorcier dans la "création d'indices".

Afin d'exercer votre sagacité, je propose à la page suivante un document très réellement distribué lors d'une réunion du groupe responsable des Lycées pour toute une académie,en 85-86. Un chiffromane impressionnant avait préparé le document "Suivi d'une cohorte", puis une suite de définitions d'indices dont je ne donne ici que les quatre premiers.

SUIVI D'UNE COHORTE. ETAT NUMERIQUE

Etablissement:..... Sept....198 Juin 198:

DATES	EVENEMENTS	Restent dans ou rejoignent leflux	Sortent du flux	Referents
Sept. X	Nbre d'élèves de 2° (y compris 2blants)	n 1		Situat;réell. [ ] a
Juin X+1	Passage en 1° dans l'établissement.....	1		
	" " dans un autre établt.....	.....	.....1'	
	Doublement dans l'établissement.....	.....	..2	
	" dans un autre établisst.....	.....	.....2'	
	Réorientation en BEP.....	.....	..3	
	Départ par abandon.....	.....	..4	
Sept. X+1	Doublants de 1°.....	n 2		Situat;réelle
	Elèves de 1° issus d'unautre établt.....	n 3		
	Nbre d'élèves de 1°.....	.....	.....	....b
Juin X+2	Passage en T dans l'établissement.....	5		
	" " dans un autre établisst....	.....	.....5'	
	Doublement dans l'établissement.....	.....	...6	
	" dans un autre établisst.....	.....	.....6'	
	Départ par abandon.....	.....	...7	
Sept. X+2	Doublants de T.....	n 4		Situat. réell
	Elèves de TY issus d'un autre établisst	n5		
	Nbre d'élèves de T.....	.....	.....	c
Juin X+3	Réussite à l'examen.....	.....	...8	
	Doublement dans l'établissement.....	.....	...9	
	" dans un autre établisst.....	.....	.....9'	
	Abandon après échec.....	.....	..10	

$R = [\ ] = n_1 + n_2 + n_3 + n_4 + n_5$

TOTAL ..... [ ]

INDICES GLOBAUX

Référent R /

		Valeur département.	Valeur académique
T 1 / taux de réussite aux examens (par rapport à l'ensemble de la cohorte)	$\frac{8}{R}$		
T 2 : taux global de passage en classe supér.	$\frac{(1+1')+(5+5')}{R}$		
T 3 : taux global de doublement	$\frac{(2+2')+(6+6')}{R}$		
T 4 : taux global d'abahdon	$\frac{4+7+10}{R}$		

On pourra observer que tous les élèves faisant leur scolarité au lycée en trois ans sont comptés à la fois dans (1) et dans (5). Le taux T2 a donc des chances d'être supérieur à 100 %...

Mais peut-être avez-vous eu déja du mal a bien sentir ce que représentait R ? En fait il s'agit de tous les élèves qui se présenteront au bac en juin X+3.

Par contre les élèves qui doublent en seconde en juin X+1 ne peuvent être ceux qui doublent leur première en juin X+2,puisque dans le meilleur des cas (doublement bénéfique!) ils entrent seulement en première à cette date...L'indice T3 ne dépasse donc pas 1.

Ces indices furent abandonnés pour cause de malformation chronique...

Un enregistrement de cette séance aurait été très exemplaire de trois attitudes possibles face aux chiffres:

- le refus viscéral de toute prise en compte d'une information numérique,

- l'acceptation béate de toutes les propositions d'indices, cela fait scientifique...

- la créativité débordante de celui qui les inventait. Reconnaissons lui le mérite d'avoir été le seul à prendre des risques.

je trouve amusant de signaler qu'ayant déposé une proposition d'animation de stage " Formation numérique et graphique en vue du suivi d'un établissement scolaire", stage réservé aux responsables et adjoints des lycées et collèges (une population d'environ 600 personnes dans notre académie), 6 seulement s'inscrivirent...et le stage n'eut donc pas lieu.

Autrement dit ,très peu de monde pense que le langage numérique nécessite une formation....Nous ne sommes donc pas encore prêts à considèrer

- qu'il s'agit d'un savoir spécifique,
- qu'il serait bon de l'intègrer de plain-pied aux savoirs scolaires.

- le pourcentage synchronique doit pouvoir atteindre 1.

C'est à dire que dans un cas extrême (dont la probabilité de réalisation est faible pour ne pas dire nulle), le nombre du dessus doit pouvoir être égal au nombre du dessous: le sous-ensemble étudié peut se révéler être le référentiel entier. Il ne doit pas y avoir d'impossibilité "congénitale" à cela.

Prenons un exemple dont la banalité vous étonnera peut-être: le taux de redoublement en seconde par exemple.

Regardons comment ce taux est traditionnellement calculé:

taux usuel t : $\frac{\text{nombre d'élèves qui redoublent en fin de seconde}}{\text{nombre total d'élèves de secondes du lycée.}}$

Or lorsque l'on prend l'effectif total des secondes, il y a parmi eux des élèves qui redoublaient déjà et donc qui ne peuvent être admis à redoubler une nouvelle fois à la fin de l'année en cours.

Supposons qu'un lycée annonce un taux régulier de 50% de redoublement en seconde. Quelle réalité cela recouvre-t-il ?

Les classes comprenant en début d'année moitié de doublants et moitié d'élèves nouveaux, en fin d'année les déjà-doublants passent et les autres s'apprêtent à doubler à leur tour. Le taux traditionnel sera bien de 50% et pourtant tous les élèves passent deux ans en seconde.

Ce taux n'obéit donc pas à la règle précédemment énoncée.

Le taux réel de redoublement devrait être calculé comme suit

Taux réel : $\frac{\text{nombre d'élèves qui redoublent en fin de seconde}}{\text{nombre total d'élèves faisant leur première seconde.}}$

Heureusement, l'écart entre le taux usuel t et le taux réel T diminue quand la situation est moins caricaturale. Ainsi,

pour un taux annoncé de 25%, le taux réel est de 33%

pour un taux annoncé de 20%, le taux réel est de 25%,

pour un taux annoncé de 15%, le taux réel est de 17,6%

enfin pour 10 %, le taux réel est de 11,1 %.

(on voit qu'il y a accélération de l'erreur quand le taux augmente.)

*** 2. Un pourcentage sert à communiquer.

J'ai en quelque sorte défini le pourcentage comme l'écriture journalistique de rapports obéissants à certaines règles.

Je n'ai pas énoncé ces règles en tant que professeur de mathématiques, mais en essayant de dégager les codes trop souvent implicites de la communication par les pourcentages.

Rien n'empêche de définir, au sens de définition mathématique, le taux de redoublement tel qu'il est usuellement calculé. IL n'empêche que si un pourcentage énoncé transgresse les règles 2 ou 3, il induira un malentendu chez ceux qui le recevront. Le repérage aux "cent pour cent" (on ne peut pas les dépasser, mais on peut les atteindre) me semble une règle "en acte" comme disent les didacticiens. Il faut en tenir compte.

• Exemple d'envie de transgresser la règle 1)

Afin de familiariser les élèves aux tableaux à double entrée je distribuai un jour le petit tableau suivant.

Répartition de la population française ayant entre 30 et 34 ans au recensement de 1975:

	Hommes	Femmes	Total
Célibataires	249 200	154 120	403 320
Mariés	1307 815	1251 835	2559 650
Veufs	3 090	11 395	14 485
Divorcés	34 690	48 430	83 120
TOTAL	1594 795	1465 780	3060 575

Tableau extrait de:

"Comprendre l'information économique et sociale."

de ML Lévy, S Ewenczyk et R Jammes, chez Hatier.

Plusieurs élèves (de seconde) proposèrent alors

$$\frac{\text{effectif des hommes célibataires}}{\text{effectif des femmes célibataires}}$$

J'usais alors de l'autorité du "prof" pour refuser que l'on exprimât ce résultat "en pourcentage", mais sans réel argument...et un élève têtu partit en maugréant qu'il ne voyait toujours pas pourquoi on ne pouvait pas mettre "les hommes sur les femmes" (!)

Pourquoi pas en effet ? Parceque les profs de math posent la première règle en guise de définition, donc d'autorité pure? Mais le mathématicien est-il, de droit, le seul à pouvoir légiférer ici?

Reprenons le travail donné aux élèves. A partir d'un tableau de données brutes, il fallait communiquer d'une quelconque façon l'ECART entre les hommes et les femmes.

- l'écart absolu, c'est-à-dire la différence, est vite laissée de côté car 9000 de différence, ce n'est pas du tout pareil si le nombre est de l'ordre du million ou de dix mille.

- mais lorsque le caractère étudié est dichotomique (noir/blanc ; ici homme/femme) il est naturel de penser au rapport des deux effectifs. D'ailleurs cela se fait. Pour simplifier, s'il y avait dans le tableau 3000 veufs et 12000 veuves, ne dirait-on pas qu'il y a

"un veuf POUR quatre veuves " ?

- cela est d'ailleurs peut-être plus parlant que le passage à l'effectif total, qui donnerait, compte tenu des priorités de la grammaire:

" quatre veufs SUR cinq...sont des veuves"

Il s'agissait donc de faire accepter à l'élève la règle 1, (laquelle contient d'ailleurs les deux autres puisque le nombre d'éléments du sous-ensemble est toujours inférieur ou égal au nombre d'élé-

ments de l'ensemble de référence).

Exercice : Une population compte 100 personnes. On note x le nombre de femmes, donc 100-x le nombre d'hommes. Etudier les variations de

$$r(x) = \frac{100 - x}{x} \quad \text{pour} \quad x \text{ variant de } 0 \text{ à } 100.$$

(On peut utiliser les connaissances sur l'hyperbole si elles ont déja été travaillées, ou bien au contraire utiliser cette fonction comme une approche de l'hyperbole, en la construisant par points successifs)

nombre de femmes	0	10	25	50	75	90	100
Rapport: Hommes/femmes	erreur	9	3	1	$\frac{1}{3}$	$\frac{1}{9}$	0

Alors le refus d'exprimer ce rapport en pourcentage devient unanime. On voit aussi se développer un argument de "symétrie": quand le nombre de femmes varie de 0 à 50, celui des hommes varie de 50 à 100 et réciproquement. Alors il parait incongru que r varie tantôt sur l'intervalle $]0,1]$ et tantôt sur $[1,+\infty[$.

En conclusion, on acceptera donc la comparaison directe de deux parties complémentaires (du genre hommes/femmes), mais cette comparaison ne s'écrit pas en pourcentage.

On pourra, de plus, faire dresser la liste de correspondance

UN pour DEUX signifie 33%

UN pour TROIS signifie 25% etc...

... Parenthèse : prenons le temps de nous arrêter sur les petits mots de la langue française qui sont liés à des opérations...

SUR est en général lié à la division, pourtant on dit qu'un terrain mesure 30 sur 50, ou qu'on va acheter une planche de 25 sur 80, alors que dans ces deux cas la division n'a aucun sens, mais plutôt la multiplication. On dit aussi 25 PAR 80

Le double DE, le triple DE évoque la multiplication, (ce sont surtout les mots double et triple...) pourtant 33% DE... ou un quart DE... devrait aussi attirer l'idée de produit. Faisons remarquer aux élèves que la structure grammaticale est la même...

Le mot POUR est moins employé.

*** 3. Importance du référentiel.

Il est évident qu'il faut faire sentir aux élèves l'importance du référentiel et que des erreurs du genre de celle faite par le journal Libération à propos de l'hystérie, sont très utiles pédagogiquement parlant!

Cela permet aussi d'habituer l'élève à être actif face à un texte imprimé,ce qui n'est pas une démarche naturelle, même pour nous!

- a) Tableau à double entrée.

Une façon de faire percevoir l'importance du référentiel est de poser des questions sur un tableau à double entrée en données brutes.

Reprenons le tableau déja cité:

	Hommes	Femmes	Total
Célibataires	249 200	154 120	403 320
Mariés	1307 815	1251 835	2559 650
Veufs	3 090	11 395	14 485
Divorcés	34 690	48 430	83 120
TOTAL	1594 795	1465 780	3060 575

- Parmi les hommes, quel est le pourcentage de célibataires?
- Parmi les célibataires, quel est le pourcentage d'hommes?

Un objectif raisonnable en seconde serait que les élèves perçoivent bien qu'à UN tableau de données brutes peuvent correspondre DEUX tableaux de pourcentages.

	Hommes	Femmes	Ensemble
Célibataires	15,6	10,5	13,2
Mariés	82,0	85,4	83,6
Veufs	0,2	0,8	0,5
Divorcés	2,2	3,3	2,7
Total	100	100	100

	Hommes	Femmes	Total
Célibataires	61,8	38,2	100
Mariés	51,1	48,9	100
Veufs	21,3	78,7	100
Divorcés	41,7	58,3	100
Ensemble	52,1	47,9	100

Tableaux extraits de:

"Comprendre l'information économique et sociale."

de ML Lévy, S Ewenczyk et R Jammes,chez Hatier.

Les référentiels seront déduits du sexe (ensemble des hommes, ensemble des femmes) ou bien de l'état matrimonial (ensemble des célibataires, ensemble des veufs,etc..). La colonne ou la ligne ne contenant que des 100 sert précisément à guider l'interprétation du tableau de pourcentages, ce sont elles qui signalent les référentiels auxquels se rattachent les pourcentages. Si ces lignes de "100" ne sont pas données avec le tableau, la première approche de lecture consiste justement à les reconstituer...

- b) Les sous-ensembles "flous".

L'ensemble de ce paragraphe est directement inspiré de la revue de l'INSEE, "Données Sociales - 1981 ". Les citations sont extraites de l'article LE CHOMAGE, page 109 à 130.

Nous avons dit qu'un pourcentage est l'écriture d'un rapport dont les deux termes dénombrent les éléments d'un ensemble (le référentiel et un sous-ensemble de celui-ci). Encore faut-il insister pour bien faire comprendre qu'en sciences humaines les ensembles ne sont pas ceux des mathématiques, et qu'il n'est pas toujours aisé de définir un ensemble...

Prenons ainsi le terme de chômeur. Un lecteur naïf, c'est-à-dire un consommateur ordinaire de journaux (ce que nous sommes tous..) se contente d'associer au terme chômeur l'idée de non travail, puis nous ingurgitons un chiffre, ou l'énoncé de ses variations du genre o,2%...

Or plusieurs "définitions" du chômeur coexistent...

* la définition du Bureau International du Travail (BIT) qui devrait permettre ,comme son nom l'indique, les comparaisons internationales...

* les "Demandeurs d'Emploi en Fin de Mois (DEFM), c'est-à dire," toute personne sans emploi, inscrite à l'ANPE comme recherchant un emploi durable à temps plein pour lequel elle est immédiatement disponible."

* les définitions utilisées dans les recensements, et qui distinguent:

- la Population Disponible à la Recherche d'un Emploi (PDRE)

- la Population Marginale Disponible à la Recherche d'un Emploi (PMDRE)

la distinction entre les deux populations étant dans l'activité déployée....pour chercher un emploi! (Par exemple n'avoir fait aucune démarche pour trouver un emploi durant le mois précédent l'enquête , vous classe dans la population "marginale").

Chacune de ces définitions à ses avantages et ses inconvénients. Les DEFM sont comptabilisées tous les mois...mais tous les "sans travail" ne font pas la démarche administrative d'aller s'inscrire à l'ANPE. Et puis le droit de s'inscrire à l'ANPE dépend de textes administratifs et législatifs qui varient souvent...Quel est alors le sens du suivi de cet indice avec des dixièmes de pour cent?

Les enquêtes d'emplois ,elles, ne dépendent pas des textes administratifs....mais de la formulation du questionnaire qui n'est pas toujours le même.... D'où un flux d'échange entre PDRE et PMDRE!

Enfin ,pour avoir une idée de comparaison internationale, voici la recette:

La définition du chômage donnée par le BIT et les recommandations annexes ont fait l'objet d'une traduction en critères factuels retenus pour classer une personne en situation de chômage. Ces critères permettent, partant de l'ensemble des personnes qui composent la PDRE et la PMDRE à l'enquête emploi, de mesurer le chômage au sens du BIT.

Ainsi, en mars 1980, PDRE et PMDRE représentaient ensemble 1 718 000 personnes. Pour parvenir au chômage BIT, il faut (tableau 63) :

— retrancher 388 000 personnes qui ne sont pas immédiatement disponibles (180 000), ou n'ont accompli aucun acte de recherche pendant le mois précédant l'enquête (193 000), ou sont à la recherche d'un emploi non salarié, ou qui ont travaillé pendant la semaine de référence ;

— ajouter 61 000 personnes dont l'essentiel est constitué par des personnes ayant trouvé un emploi qui commence ultérieurement et celles qui n'ont pas travaillé pendant la semaine de référence en raison de la mauvaise conjoncture économique.

On obtient un chiffre de 1 391 000 chômeurs au sens du BIT (83 % de la PDRE et 58 % de la PMDRE appartiennent au chômage BIT).

Nous n'insterons pas davantage sur la difficulté de bien cerner l'appartenance ou la non-appartenance à un ensemble. Il s'agit seulement d'en prendre conscience afin que la précision des chiffres donnés ne serve pas d'écran magique...

---

* EXERCICE paru dans LE MONDE DE L'EDUCATION. Mars 87 .

Que penser du taux de réussite par tranches?

TAUX DE RÉUSSITE PAR TRANCHES

I : moins de 40 %	VI : de 65 à 69 %
II : de 40 à 49 %	VII : de 70 à 74 %
III : de 50 à 54 %	VIII : de 75 à 79 %
IV : de 55 à 59 %	IX : de 80 à 89 %
V : de 60 à 64 %	X : plus de 90 %

Attention, lorsque le « rendement pédagogique » d'un lycée (c'est-à-dire le rapport reçus au bac/nombre d'inscrits en seconde trois ans auparavant — la quatrième ligne des tableaux) se situe dans la tranche X, cela peut signifier que cet établissement amène tous ses élèves de seconde au bac en trois ans, ce qui est rare : il est alors excellent. Mais cela peut signifier aussi que, pour diverses raisons, il a accueilli de nouveaux élèves en première ou en terminale. Le nombre d'élèves de terminale se trouve dès lors supérieur à celui des secondes trois ans auparavant, élevant artificiellement le taux de rendement du lycée. Conclusion : renseignez-vous auprès du proviseur ou auprès des associations de parents.

Les élèves passant le bac en quatre ans ou cinq ans ne figurent pas non plus dans les "inscrits en seconde trois ans auparavant" et ils sont cependant comptés dans le numérateur. Plus il y a d'élèves faisant le cursus en 4 ou 5 ans, plus le taux de rendement est artificiellement élevé... Ces élèves ajoutés clandestinement au numérateur ne sont pas en quantité négligeable : plus du tiers des reçus ont passé 4 ou 5 ans au lycée...

Il vaudrait mieux écrire ce rapport sous forme décimale, comme un indice. La forme en pourcentage est trompeuse.

*** 4. Comparaison de pourcentages.

Comparer des pourcentages portant sur un même référentiel offre peu d'intérêt puisqu'on pourrait comparer directement sur les données brutes. Cela permet tout au plus d'énoncer des chiffres plus simples donc peut-être plus aisés à retenir: sur l'ensemble de la population des 30-34 ans, il y a 13 % de célibataires et 83% de mariés; "environ six fois plus" se perçoit mieux que sur les chiffres absolus.

Il est par contre indispensable d'utiliser des pourcentages pour comparer des rapports "isomorphes".

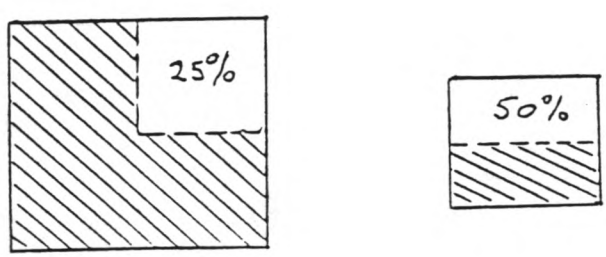
Expliquons ce qualificatif sur un exemple:

Y a-t-il plus d'hommes célibataires parmi les hommes que de femmes célibataires parmi les femmes (toujours entre 30-34ans)? Autrement dit, on est face à deux structures parallèles et on voudrait comparer des rapports. Le pourcentage réalise de fait le "dénominateur commun" nécéssaire pour comparer deux fractions. Bien sûr, si les deux référentiels ont un effectif assez voisin, ce qui est le cas dans notre exemple, les données brutes permettent de faire la comparaison approximative directement. Mais lorsque les échelles sont très différentes, le pourcentage est indispensable pour comparer des structures.

Ainsi, il ne sert à rien de comparer les chiffres absolus des budjets de l'Education Nationale de la France et de l'Algérie. Par contre la part de l'Education Nationale dans le budjet global de chaque pays montre l'investissement fait par les pays en voie de développement dans la scolarisation de leur jeunesse.

A l'inverse, il faut bien insister, car c'est une erreur fréquente dans les devoirs de géographie et d'économie, que la comparaison des pourcentages ne renseigne pas sur l'ordre des chiffres absolus:

25% d'un "gros gateau" peut représenter plus que 50% d'un petit, en données absolues s'entend.



*** 5 . Addition de pourcentages.

- On ne peut additionner que des pourcentages portant sur un MEME référentiel, est la première condition élémentaire à respecter.

- L'addition de deux pourcentages ne doit pas amener à compter deux fois les mêmes éléments.

Prenons un exemple.En admettant que 20% de la population de Grenoble lit le journal "Le 38", et que 15 % de cette même population lit "Grenoble - Info", si on met une petite annonce dans chacun de ces deux hebdomadaires, on ne saurait en déduire que l'on peut atteindre ainsi 20 + 15 soit 35% de la population grenobloise, puisque,sans aucun doute, certaines personnes lisent ces deux journaux...

On rencontre ici une loi générale pour toutes les "mesures". On ne peut ajouter des quantités qui se "recoupent" pourrait-on dire, pour parler simplement.

Cela paraît simple en effet....Mais en pratique, lorsqu'on se retrouve devant un tableau numérique, il n'est pas toujours aisé d'éviter les pièges...

En guise d'exercice, nous proposons un tableau des résultats au bac concernant l'académie de Grenoble (Publication du Rectorat):

(Voir pages suivantes)

- Où l'on retrouve la moyenne pondèrée.

Prenons des chiffres simplifiés pour expliquer.

15000 élèves entrent ensemble en seconde. Trois ans plus tard, 40% sont candidats à un bac général et 20% à un bac Industriel ou Tertiaire. Parmi ces candidats, la réussite au bac général est de 80%, celle aux bacs TT+TI de 60%.

La moyenne "arithmétique" (qui donnerait (80 + 60) : 2 = 70) conduirait à faire additionner des pourcentages de référentiels différents. Mais que fait-on quand on pondère? C'est à dire lorsqu'on fait compter "deux fois plus" la réussite au bac général puisque les candidat sont deux fois plus nombreux?

En fait lorsqu'on pose le produit 0,80 . 0,40 (c'est-à dire 80% de réussite parmi les 40% des 15000 élèves, on se ramène au référentiel initial de la cohorte totale entrée en seconde. On fait de même en pondérant 0,60 . 0,20. Ces deux produits donnent deux pourcentages du MEME référentiel 15000, donc l'addition est légitime... Le troisième chiffre de la dernière colonne serait,pour notre exemple:

(0,80 . 0,40) + (0,60 . 0,20) = 0,44

Dans le tableau réel : (0,746 . 0,393)+(0,679 . 0,186) = 0,419 soit 42% Sur les 13852 élèves entrés en seconde en 82, 42% ont un bac 3 ans plus tard.

EVALUATION DES RESULTATS
OBTENUS EN 1985

EXERCICES . (SUR DES DONNEES REELLES)

RECTORAT DE L'ACADEMIE DE GRENOBLE

Voici des chiffres concernant les 13852 élèves entrés en seconde en septembre 82 dans l'académie de GRENOBLE.

ACADEMIE DE GRENOBLE — LE : 17/03/86

ECHL10 — C U R S U S S C O L A I R E D E S L Y C E E N S — PAGE : 00019

C O H O R T E 1982 - 1984 S U R 3 A N S

	ELEVES SECONDE	ELEVES DE TERMINALE					SUCCES AUX BAC/BTN				
		GENER.	T.INDUS.	T.TERT.	TI + TT	TOTAL	GENER.	T.INDUS.	T.TERT.	TI + TT	TOTAL
TOTAL ACADEMIE	13852	5444	1014	1575	2589	8033	4065	710	1050	1760	5825
		39,3	7,3	11,3	18,6	57,9	74,6	70,0	66,6	67,5	72,5
TAUX REUSSITE/ENTREE							29,3	5,1	7,5	12,7	42,0
	F 7790	F 3384	F 114	F 1215	F 1329	F 4713	F 2549	F 77	F 819	F 896	F 3445
	G 6062	G 2060	G 900	G 360	G 1260	G 3320	G 1516	G 633	G 231	G 864	G 2380

- Quels sont les référentiels auxquels se rapportent 39,3%? 18,6%? 74,6%? 29,3%? 42,0%?

- On sent bien qu'on ne peut ajouter 74,6 et 67,9 parce que "ça dépasserait 100%", mais on peut être tenté de faire la moyenne arithmétique, ce qui inclut tout de même l'addition. Comment peut-on retrouver le 42,0% à partir de 74,6 et de 67,9 (et d'autres colonnes)?

La page précédente répond à la dernière question qui a pour but de comprendre la notion de moyenne. La page suivante donne le "mode d'emploi" pour "consommateur de résultats", mode d'emploi extrait lui aussi de la brochure du Rectorat. Mais en fait, pour bien "entrer" dans un tel tableau, il faut prendre une calculette et essayer de retrouver soi-même les pourcentages à partir des chiffres bruts qui sont donnés.

TABLEAUX NUMERIQUES : MODE D'EMPLOI

* Pour la rubrique "Elèves de Terminales":

Tous les pourcentages sont calculés à partir du même référentiel, à savoir le nombre total d'élèves entrés en secondes. De ce fait, les pourcentages s'additionnent naturellement,comme les données. La colonne "total" est la somme des colonnes 1 et 4 ,"general" et "Ter industrielles + Term tertiaires "

(On pourra parfois constater quelque différence portant sur la décimale,cela vient des "arrondis" qui sont faits "grossièrement",c'est à dire systématiquement à la valeur inférieure:
1760 : 2589 donnant 67,98 % est arrondi à 67,9 % ....)

* Pour la rubrique "Succès au BAC / BTN " :

Nous voyons apparaitre ,en dessous des données, deux lignes de pourcentages.

- la seconde ligne se calcule comme pour la rubrique précédente, à partir du même ensemble de référence (le nombre total d'élèves de secondes), donc ici aussi les pourcentages s'ajoutent.

- mais pour la première ligne,il s'agit de pourcentages "semblables" ou encore pourrait - on dire, "calculés en parallèles", c'est-à-dire chacun par rapport à son propre référentiel: Chaque résultat d' une colonne SUCCES est divisé par la colonne correspondante de la rubrique ELEVES DE TERMINALES :

$$\frac{\text{nombre d'élèves reçus au bac général}}{\text{nombre d'élèves de la cohorte en term. Générale}} \quad ; \quad \frac{\text{nombre d'élèves reçus au bac industriel}}{\text{nombre d'élèves de la cohorte en Term Industrielle}}$$

Ces pourcentages ressemblent aux rituels taux de réussite au bac,mais ils ne prennent en compte que les élèves faisant leur première terminale puisqu'il s'agit ici du suivi d'élèves entrés la même année en seconde.)

Ces pourcentages ne s'ajoutent pas, (en tous cas pas de manière "arithmétique", il faudrait faire intervenir des moyennes pondérées).

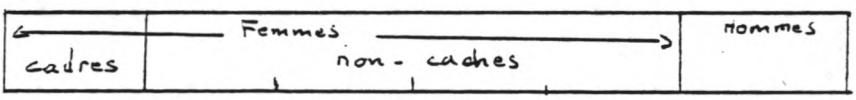
* Afin de ne pas épuiser le lecteur, aucun pourcentage n'est donné en ce qui concerne les répartitions par sexe des lycéens. Les mêmes calculs que précédemment peuvent être faits ,en parallèle , pour filles et garçons.

*** 6. Produit de pourcentages synchroniques.

On rencontre ici l'une des grosses difficultés de nos élèves En fait le produit de pourcentages vient de la considération d'un sous-ensemble d'un ensemble, lequel est lui-même partie d'un ensemble plus vaste. C'est le changement de référentiel qui gène nos élèves.

Exemple: Dans une usine de textile, 80% des employés sont des femmes. Et parmi les femmes, 20% sont cadres (cadre moyen ou supérieur). Parmi le personnel, quel est le pourcentage de femme-cadre?

Bien sûr, j'ai fait des petits dessins...mais il n'empêche que beaucoup d'élèves ne sentent pas pourquoi on multiplie, car la multiplication par un nombre inférieur à 1 est, d'une certaine façon, contre nature. Peut-être est-ce justement par le maniement des pourcentages que l'on peut aider les élèves à l'apprivoiser?

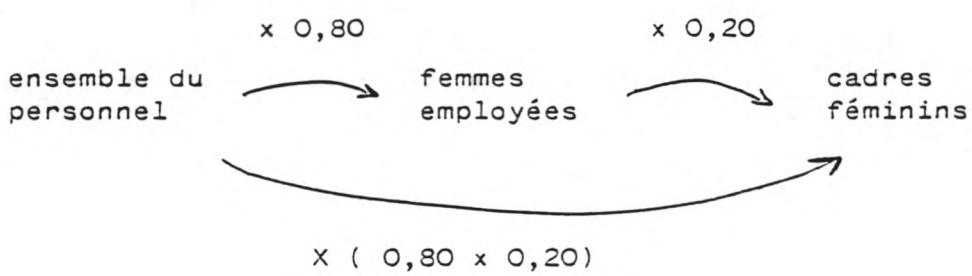


Avec les pourcentages de pourcentages on rejoint un domaine à priori bien différent, celui de la composition des fonctions. Le changement de référentiel, c'est le changement d'ensemble de définition.

Comme j'écrirais "le double du carré de" en décomposant les génitifs:

(en remarquant l'inversion de l'ordre, propriété des génitifs dans la langue française et non décision arbitraire des profs de math dans la notation de la composée de deux fonctions),

j'écrirais de même:



Le chemin direct, la composée, montre donc que les cadres féminins représentent 16% du personnel de l'entreprise.

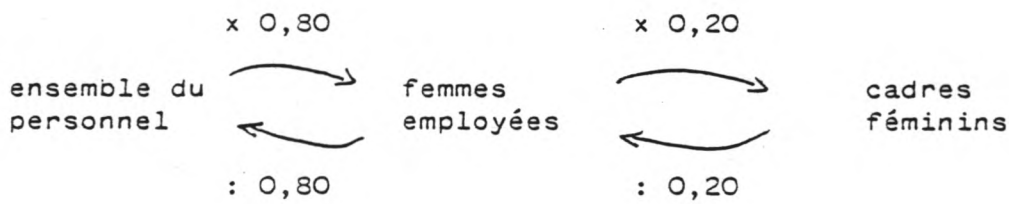
Il me semble que les choses passent mieux

- en travaillant d'abord sur les composées de fonctions, ce qui permet une relecture en profondeur de l'écriture algébrique. L'usage des calculettes programmables est ici très bénéfique : la décomposition prépare le programme.

- puis sur les réciproques simples : ce qui "défait" ajouter 3, c'est soustraire 3, ce qui défait " mulplier par 4" c'est diviser par 4.
(D'où la récupération de certaines erreurs systématiques :

$3a = 15$ alors $a = \frac{15}{-3}$ car "on change de membre, donc on change de signe...)

- Alors si on décompose les pourcentages de la même façon, les élèves sauront mieux "remonter", c'est-à-dire répondre à une seconde question de l'exercice : Sachant qu'il y a 4 femmes cadres, combien y a-t-il d'employés dans l'entreprise?



Soit, en lisant de droite à gauche :

PARENTHESE POUR ENSEIGNANTS DE MATH EN LYCEE

La composée me semble le carrefour, ou le nœud, des difficultés de nos élèves dans des domaines qui semblent à première vue éloignés de ce sujet.

-une élève de Terminale D n'arrive pas à dériver $\ln \frac{2x+1}{x-2}$ alors qu'elle "sait les formules par coeur"...
Oui, mais elle ne sait pas dans quel ordre utiliser les dites formules et commence par la formule du quotient, puis multiplie par 2 / 2x+1 pour placer la formule u'/u !

- un autre ne sait pas trouver les limites, car il ne sait pas utiliser ses connaissances de cours dans le bon ordre. Il ne distingue pas la fonction donnée de

$$\frac{(2 \ln x) + 1}{(\ln x) - 2}$$

- De même pour l'ensemble de définition....

Autrement dit le travail sur les composées me semble fondamental, et il n'est pas nécéssaire de connaître le "cours" pour travailler dessus, au contraire serais-je presque tentée d'écrire. Faisons fonctionner des composées chaque fois que nous en rencontrons.

*** 7. Ecriture décimale d'un pourcentage synchronique.

Il me paraît important de bien distinguer l'écriture usuelle du pourcentage et son écriture décimale.

Beaucoup d'élèves arrivent en seconde avec des difficultés sur le calcul des pourcentages. Il semble que le mode de calcul soit pour eux "quelque chose à apprendre", un effort de mémoire à l'état pur. Ainsi, il est fréquent d'entendre:

- on divise par cent ?
- non, il faut multiplier par 100, rétorque un autre...

Une fois de plus, prenons un exemple. Si nous avons 420 élèves reçus sur 640 candidats présentés, le rapport de la partie au tout a du sens:

..elle amène à poser $\frac{420}{640}$

.. mais quel peut bien être le SENS du "multiplié par 100" ?

Pire encore est le non-sens si on commence par la multiplication, ce que permettent les règles opératoires:

420 reçus x 100.....? le tout ensuite divisé par 640!

Je fais l'hypothèse que ce manque de sens gêne une bonne partie de nos élèves (mais pas tous! Il y a même chez les "bons en math" scolairement parlant, une grande propension à ingurgiter du "vide de sens". Peut-être ont-ils, de cette façon, l'impression d'atteindre les sommets enviés de l'abstraction pure ? Peut- être que notre manière de penser, de montrer , de dire , d'enseigner les mathématiques induit cette hiérarchie? Mais je fais aussi l'hypothèse que ces "avaleurs de formules " resteront au stade de l'exécution et ne pourront devenir des mathématiciens qu'en donnant du sens à ce qu'ils font...)

Quant aux élèves qui réclament que les calculs aient du sens, faute de quoi ils les refusent, ils ne sont pas gênés d'admettre comme une convention d'écriture qu'un rapport "d'une partie au tout" peut se traduire

- sous forme décimale pour les calculs,
- sous forme "pour cent" pour communiquer.

Pour l'addition, le calcul sous forme "pour cent" n'est pas gènant, puisque toutes les fractions ont le même dénominateur. Il peut donc cohabiter avec le calcul en nombres décimaux.

C'est avec le produit que naissent les plus grosses difficultés.

En effet, l'énoncé d'un pourcentage s'accompagne en quelque sorte d'une "unité": il y a 62% de reçus... (à vrai dire il faudrait penser reçus/ candidats...). S'il est légitime d'ajouter des "reçus", on ne voit pas bien, par contre, ce que peut signifier un produit...

L'écriture décimale permet mieux de dégager l'aspect "opérateur". On reprend ensuite la formulation en pourcentage pour rédiger, communiquer une conclusion.

Peut-être faut-il tenir compte des capacités que nos élèves acquièrent du fait de leur environnement, qui n'était pas le nôtre. Ils sont souvent à l'aise dans la communication et changer de niveau de langage est assez banal pour beaucoup d'entre eux...

C'est ainsi que voulant appliquer l'idée exposée par P. Meirieu dans son livre " L'Ecole, mode d'emploi" (trouver des activités qui associent en même temps ce que l'élève maîtrise bien et ce qui le gène, pour que l'un fasse avancer l'autre), j'en arrivai à demander,en devoir de math, un article de journal...ou une lettre à un inspecteur d'académie!

(je précise que j'accepte que des élèves rendent un devoir de math traditionnel pour m'expliquer la moyenne pondérée,s'ils préfèrent...Mon seul objectif est d'amener le plus d'élèves possible à s'intéresser au concept de math en question, et de ne pas détourner à priori , des élèves créatifs du domaine des mathématiques: elles en ont besoin! On dit trop souvent que les jeunes ont besoin des mathématiques, que ne pense-t-on aussi à la réciproque...L'actuelle pénurie d'étudiants en math est la conséquence directe de cet "oubli" dans la tète des enseignants du secondaire. Refermons cette parenthèse pour retourner sagement à nos pourcentages... Je me suis aussi donné un objectif : pas plus de 8 jours (pleins!) pour mettre sur papier cette seconde partie du stage Math-Géo! On exige bien de nos élèves qu'ils travaillent en temps limité ...)

... II ... A PROPOS DU POURCENTAGE DIACHRONIQUE.

*** Ecriture décimale d'une "augmentation" ou d'une "baisse"

Si on note p_a le prix ancien (ou le prix initial) d'un objet,les élèves savent assez bien passer au "nouveau prix" en décomposant:

Par exemple, après une augmentation de 12%:

le nouveau prix, noté p_n $\quad p_n = p_a + p_a \cdot 0,12$

Et s'il s'agit d'une baisse de 12%:

$$p_n = p_a - p_a \cdot 0,12$$

Un objectif accessible en seconde est d'imposer la factorisation du prix ancien pour obtenir directement le nouveau prix. On peut alors observer le blocage des élèves à une factorisation ne portant que sur des nombres....Ils ne savent factoriser qu'avec des lettres!

En effet , en classe, avant de noter avec une lettre le prix initial et le nouveau prix, nous faisons beaucoup d'exemples numériques, et on sent un refus de factoriser 352 dans l'écriture:

$$352 + 352 \cdot 0,12$$

(On rencontre aussi l'obstacle de 352 qui est 352 . 1)

Il paraît pourtant commode d'arriver à une bonne maîtrise du résultat suivant :

Augmenter de 12% revient à multiplier par 1,12
Baisser de 12% revient à multiplier par 0,88

On retrouve ici la difficulté du passage à une formule non signifiante pour l'élève. Encore le 1,12 montre bien qu'il est le résultat de 1 + 0,12 , "1" pour le prix ancien, et 0,12 pour l'augmentation, mais le 0,88 cache bien son jeu... et il faut patiemment reprendre l'explication, apprendre à voir dans 0,88 : 1 - 0,12...

Quand j'écris qu'il faut reprendre l'explication, je veux dire qu'il faut accepter que les élèves demandent successivement l'explication... cela est certes un peu ingrat; pourtant un élève qui pose une question prouve qu'il vient de se livrer intérieurement à une activité mathématique, et cela ne sert à rien de lui dire " Ton voisin vient de poser la même question, tu n'avais qu'à écouter!" Ecouter peut très bien être une forme d'activité, mais cela peut aussi être totalement passif et dans ce cas, si l'élève n'a pas cheminé au même rythme que celui qui a énoncé la question, la réponse ne lui apporte rien. S'il s'intéresse à ce qui se passe dans la classe, il est

donc logique qu'il formule à son tour la question, entraînant avec lui ceux qui avaient besoin de plus de temps. Il faut s'attendre à voir de nouveau poser la question au cours suivant, pour espérer atteindre une couche plus profonde de compréhension.

Ayant expliqué aux élèves que sur certains points difficiles il est normal que la même question soit posée plusieurs fois, je peux dire que cela ne dépasse pas trois fois, exceptionnellement quatre, pendant la même séquence. Ce sont alors des élèves qui reprennent la réponse à leur façon pour l'expliquer aux autres. Un élève qui s'est échappé un moment du travail en classe en rêvant ou en discutant d'autre chose avec son voisin, réintègre en général l'activité le plus discrètement possible, et non en se faisant remarquer en posant une question! Et quand bien même cela arriverait parfois, mieux vaut que le "doute" profite à l'élève...

A propos de cet exemple, je déborde un peu sur l'attitude d'écoute en classe. Si on pense que l'activité première du temps en classe se passe intérieurement, dans les têtes "apprenantes" et non dans LA tête "enseignante", cette attitude est logique... (Reste à la mettre en oeuvre heure après heure... ce qui n'est pas toujours facile, mais le devient de plus en plus lorsqu'on a analysé une fois une telle situation.)

*** Les intérêts composés...

Le passage de la communication " augmenter de 12%" à l'opération "multiplier par 1,12" devient particulièrement rentable quand on réitère d'années en années.

On peut même profiter de ce thème porteur (est-ce l'influence de la crise économique, les jeunes n'ont jamais été aussi sensibles aux variations de prix...) pour réconcillier certains élèves avec les exposants.

Supposons qu'on se penche sur l'évolution du prix de la place de cinéma en supposant une augmentation de 5% chaque année.

Notons Po le prix initial, P1 le prix dans un an, P2 le prix dans deux ans..... P15 le prix dans 15 ans.

Chaque année le prix est donc multiplié par 1,05, ce qui donne la suite de prix suivante:

$P1 = PO \cdot 1{,}05$

$P2 = P1 \cdot 1{,}05 = PO \cdot 1{,}05 \cdot 1{,}05 = PO \cdot (1{,}05)^2$

$P3 = P2 \cdot 1{,}05$ soit $P \quad PO \cdot (1{,}05)^3$

etc...

Dans 15 ans $P15 = PO \cdot (1,05)^{15}$

A titre d'exemple, si on prend PO = 38f, cela fera P15 = 79f.

(avec une hypothèse de 6% d'augmentation pendant 15 ans, cela ferait 91 f...)

Il importe cependant de bien faire distinguer ici l'hypothèse (exemple : les prix augmentent chaque année de 5% pendant 15 ans) et l'outil mathématique qui travaille sur cette hypothèse-là.La rigueur de l'outil n'est pas le garant de la rigueur du résultat. D'ailleurs avec les petites calculettes, il est désormais très facile de faire plusieurs modèles en variant les hypothèses.

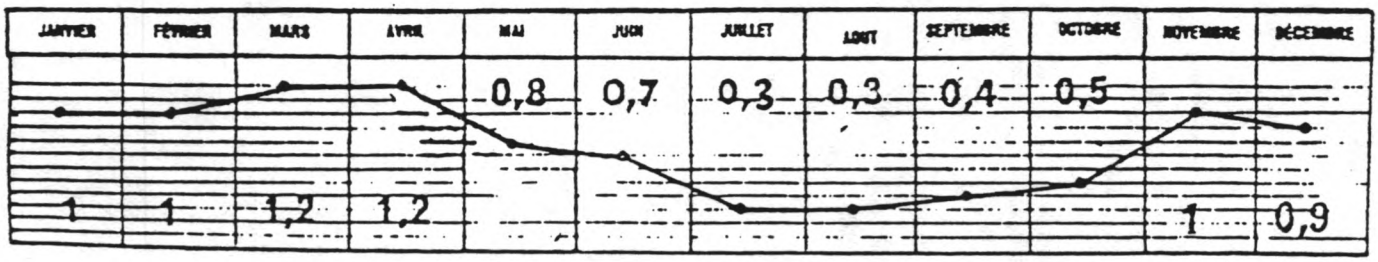
On retrouve ici encore des pourcentages qui ne s'ajoutent pas: augmenter de 8%, puis de 12% ne donne pas au total une augmentation de 20%, mais comme $1,08 \cdot 1,12 = 1,21$,une augmentation réelle de 21%.

Nous donnons page suivante un bel exemple de pourcentage que les lecteurs du Dauphiné Libéré avaient envie d'ajouter, croyant prendre leur journal en défaut. La réponse étant très clairement formulée...nous la donnons aussi!

Le Dauphiné Libéré du 13 Janvier 1983:

LIMITATION DE LA HAUSSE DES PRIX

9,7 % : OBJECTIF ATTEINT



Et sa réponse aux coups de téléphone des lecteurs, le lendemain:

INFLATION

C'est bien 9,7 %

NOMBREUX coups de fil hier matin à notre rédaction de Veurey pour contester le taux d'inflation que nous annoncions à « la une » pour l'année 1982 soit 9,7 %.

A juste titre les lecteurs, en ajoutant les taux mensuels les uns aux autres ne parvenaient pas à 9,7 %, mais à 9,3 % (1).

Certes, pas de quoi chipoter, une différence de 0,4 % ce n'est pas le Pérou ! Mais quand même.

Pourtant, il s'agissait bien de 9,7 % tout simplement parce que le calcul utilisé par nos lecteurs n'était pas le bon.

Pour connaître, en effet, le taux annuel d'inflation, il n'est pas aussi simple d'ajouter les uns aux autres les taux mensuels. On applique le premier taux à l'indice précédent et c'est sur le nouvel indice que l'on applique le taux suivant, et ainsi de suite.

Ainsi à la fin de l'année sur une base de 100 F en décembre 82, on parvient à 109,76 F en décembre 1983, soit en effet une augmentation de 9,76 % et non pas 9,3 %.

Il ne sera pas inutile de préciser que ce mode de calcul s'applique également pour les revalorisation des taux horaires, par exemple.

A titre indicatif, signalons qu'un prix à la consommation de 100 F en 1970 en France vaut, selon ce calcul, 304 F en 1982, soit une augmentation de 200 %. Presque du grand galop !

Marc VARLET

(1) L'addition des indices de 1982 soit : 1 + 1 + 1,2 + 1,2 + 0,8 + 0,7 + 0,3 + 0,3 + 0,4 + 0,5 + 1 + 0,9 = 9,3 %

En page 2

POURCENTAGE AU COURS DES ANS...

(MODE D'EMPLOI)

*1. Pour augmenter un prix de 12%, on le multiplie par 1,12

Pour baisser un prix de 7%, on le multiplie par 0,93.

*2. Pour trouver le pourcentage de variation, on divise la nouvelle valeur par l'ancienne ("le nouveau sur dessus l'ancien, comme les étiquettes de prix dans un Hypermarché !")

$$\frac{\text{Nouveau prix}}{\text{Ancien prix}} = \begin{cases} \text{si plus grand que 1 ; 1,05} \\ \text{c'est une augmentation de 5\%} \\ \text{si plus petit que 1 : 0,92} \\ \text{c'est une baisse de 8\%} \end{cases}$$

* 3. Plusieurs années à taux constant:

Si une population baisse pendant 10 ans de 2% chaque année:

Population dans 10 ans = population initiale . $(0,98)^{10}$

§ Mode de calcul sur calculette TI57 II
§
§ Soit [Prix initial] [x] [0,98] [y^x] [10] [=]
§
§ Soit, pour se rendre compte du coefficient multiplicateur
§ (qui ne dépend pas du prix étudié):
§
§ [0,98] [y^x] [10] [x] [prix initial] [=]
§
§ Soit si on veut faire le même calcul pour plusieurs prix
§ on met en mémoire le coefficient multiplicateur:
!
§ [0,98] [y^x] [10] [=] [STO_0]
§
§ Puis [Prix initial] [x] [RCL_0] [=]
§

(Avec une TI 30, la façon de procéder est la même, mais on tapera seulement STO pour mettre en mémoire, et RCL pour utiliser ce qui est dans la mémoire. Dans la TI57, comme il y a plusieurs mémoires elles sont numérotées comme on numérote spontanément les tiroirs d'une commode : c'est dans le troisième...)

* 4. Taux annuel moyen.

Si on connaît un prix initial et un prix 5 ans plus tard, le taux annuel moyen est le pourcentage constant sur les 5 années qui aurait permis de passer du même prix initial au même prix final. C'est donc le pourcentage annuel constant qui produirait le même effet que la réalité observée (mais on ne sait pas ce qui s'est réellement passé entre les 5 ans).

Or : Prix initial . $(1 + d)^5$ = prix dans 5 ans

donc $(1+d)^5 = \frac{\text{prix dans 5 ans}}{\text{prix initial}}$

Exemple : Le prix est passé de 4,60 à 5,80 en 5 ans

§

§ $\frac{5,80}{4,60} = 1,26$ donc $(1+d)^5 = 1,26$

§

§ Taper : [5,80] [:] [4,60] [=] [INV] [y^x] [5] [=]

§

§ Résultat : 1,047 ; d (écriture décimale) vaut 0,047

§

§ donc le taux annuel moyen est de 4,7% de hausse.

§

§

§ [2] [y^x] [5] [=] fait calculer ?? dans $2^5 = ??$

§

§ et [2] [INV] [y^x] [5] [=] fait calculer ?? dans $(??)^5 = 2$

CAS PARTICULIER FREQUENT.

On ne dit pas les valeurs initiales et finales, mais par exemple on sait qu'une population a doublé en 20 ans. Quel est son taux annuel moyen?

On sait que $\frac{\text{population Dans 20 ans}}{\text{population initiale}} = 2$

donc $(1+d)^{20} = 2$ soit un taux annuel moyen de 3,5 %

* Approximations commodes.

Pour des taux petits et pour de faibles durées (2,3ans) l'erreur faite en ajoutant les pourcentages est si minime...qu'on peut s'autoriser à la faire (ce qui permet un calcul de tête).

Si un prix a augmenté de 6% en 3 ans, cela correspond pratiquement à un taux annuel moyen de 2% (calculs exacts : 1,96%...)

*** 5. Si on cherche à déterminer un nombre d'années...

Exemple : en supposant une augmentation annuelle de 7% chaque année, en combien d'années un prix est-il triplé?

Le prix dans n années = prix initial . $(1,07)^n$

donc il faut trouver n pour que $1,07^n = 3$

§
§ - On peut chercher en tatonnant...
§
§ - On peut "faire tourner" sa machine si elle est programma-
§ ble. C'est-à-dire lui faire écrire les prix successifs
§ jusqu'à ce qu'on trouve le triple, en partant de n'importe
§ quel prix initial (de 1 par exemple)
§
§ programme:
§
§ [LRN] [x] [1,07] [=] [2nd|Pause] [RST] [LRN]
§
§ Puis on tape [1] [R/S].. et on guette, en comptant
§ les tours à quel moment cela dépasse 3.
§ (Ici au cours de la 17ème année)
§ Si on trouve que les nombres défilent trop vite,on
§ retape 2nd Pause avant RST.
§
§ Si on veut voir défiler les prix successifs d'un objet
§ donné, on garde le MEME programme, mais on tape par ex:
§ 38 R/S et on aura les prix futurs de la place de cinéma
§ dans l'hypothèse d'une augmentation constante de 7%.
§
§ - Enfin on peut utiliser la touche log
§ (consulter la page "histoire d'un log" ci-après)

HISTOIRE D'UN LOGARITHME...

Sur la planète Logarithme, la vie est bien plus simple que sur terre! En particulier en ce qui concerne les opérations.

Pour parler le langage des logarithmes, il suffit d'avoir un dictionnaire. Avant il était "sur papier", cela s'appelait la table de log. Maintenant, ce dictionnaire est stocké dans les calculettes ayant la touche log.

Ce qui se dit sur Terre 1986 . . . . . se dit 3,298 chez les LOG

on tape 1986 log

Ce qui se dit 1,5 chez les LOG

se dit 31,62 sur Terre

on tape 1,5 INV log

QUEL EST L'INTERET DE PARLER LOG ?

Sur TERRE	Chez les LOG
Si on multiplie.	on ajoute.
Si on divise	on soustrait,
Si on élève à la puissance 6	on multiplie par 6
Si on veut la racine cubique	on divise par 3.

Exemple. Quel est le coté d'un cube de volume 50 m^3 ?

Il faut la racine cubique de 50

On va traduire 50 1,69897

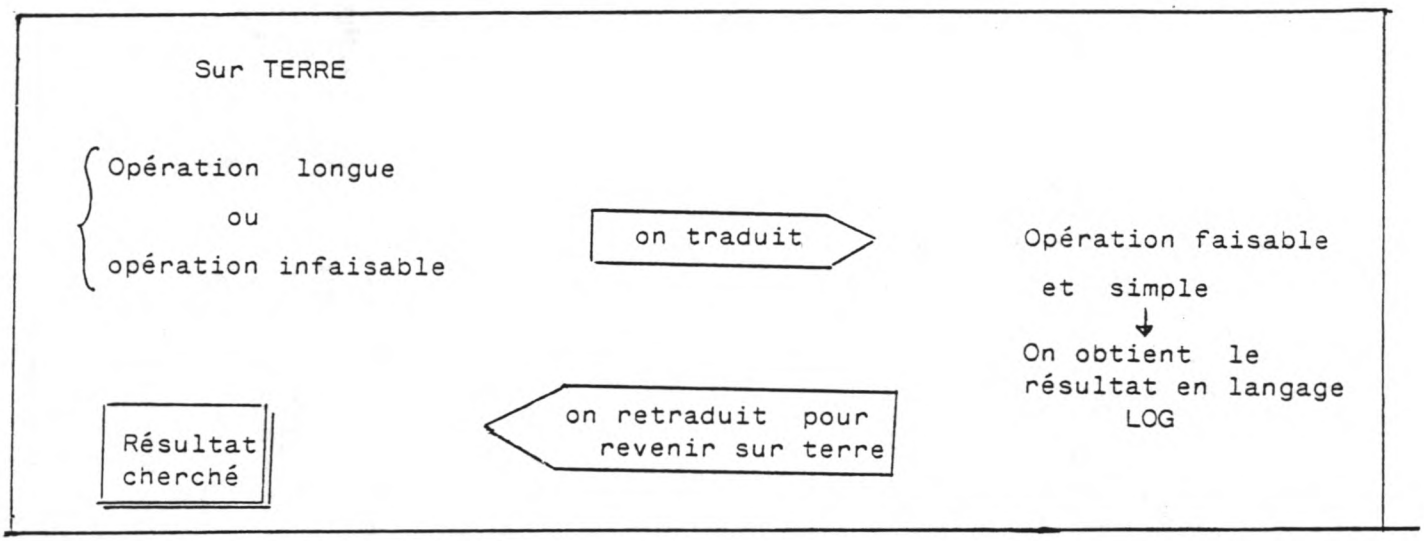
on divise par 3 :

0, 5663233

on retraduit pour retourner sur terre

3,68 m

Désormais les calculettes font cela automatiquement. Mais cela fut pendant plusieurs siècles la seule façon de faire des calculs que l'on ne savait pas faire "directement", comme extraire une racine cinquième par exemple. Ou encore pour faire des calculs qui seraient beaucoup trop longs: calculer 1,25 puissance 10 c'est long, alors que traduire en log et multiplier par 10 c'est rapide.



● Ainsi, pour calculer le taux moyen annuel d'augmentation d'une consommation qui triple en 15 ans, il fallait résoudre une équation du type

$X^{15} = 3$ qui se traduit en $15 \cdot \log X = \log 3$

$X = 1{,}076$ ⟵ $\log X = 0{,}03181$

(d'où un taux de 7,6%)

Les machines font cela toutes seules quand on tape [3] [INV] [y^x] [15] [=]

Mais dans l'exemple suivant, on utilise encore les logarithmes, car la traduction donne une phrase composite : le nombre cherché s'écrit directement en fonction de logarithmes des données, il n'y a pas à le retraduire.

Exemple. Quelle est la puissance de 1,07 qui vaut 3 ?

$(1{,}07)^n = 3$

On ne sait pas faire ⟶ On traduit ⟶ $n \cdot \log(1{,}07) = \log 3$

donc $n = \dfrac{\log 3}{\log 1{,}07}$

On trouve 16,23.

Donc $1{,}07^{16}$ est inférieur à 3, mais $1{,}07^{17}$ dépasse 3.

(Taper [3] [log] [:] [1,07] [log] [=])

III – CINQUANTE MILLIONS DE CONSOMMATEURS DE STATISTIQUES

Nous proposons ici une approche succinte de ce que sont les Statistiques, à l'usage des "consommateurs" que nous sommes tous.

Il ne s'agit surtout pas d'un cours, c'est-à-dire de ce que malheureusement on présente aux élèves de terminales, parce que c'est le programme du bac: des formules dont souvent le sens nous échappe....

Cette demande d'éclaircissement au sujet des statistiques a été formulée par des enseignants dès le premier stage Mathématiques et Géographie, mais il faut dire que nous sommes ici à la limite de nos compétences..ou au début de notre incompétence! Suite à cette demande, je partis en chasse d'un livre qui ne se contenterait pas de poser des définitions, ou de donner des résultats statistiques,mais qui parlerait surtout de la légitimité de certains calculs, du champ d'application des méthodes...

Je dois dire que j'ai été souvent découragée et rebutée: beaucoup de manuels étaient d'une lecture trop difficile pour le prof du secondaire que je suis... Ce n'est qu'au hasard d'un passage aux Presses Universitaires de France, et grâce à l'intelligence patiente d'un vendeur compétent que je dénichais enfin un livre clair, drôle, concret, rédigé en français, et où , quel réconfort, "j'ai tout compris!" Donc, bien sûr, je le recommande à ceux qui voudraient en savoir plus que la brève présentation que je vais faire ici.

Les Statistiques
Une approche nouvelle
par Sanders , Murph et ENG

Mc Graw - Hill Editeurs

Les statistiques ont un triple visage:

- elles sont de l'ordre de la COMMUNICATION et de la DESCRIPTION quand leur objectif se borne à transmettre une information contenue dans des données numériques trop nombreuses pour être saisies directement.

- elles sont de l'ordre de la PREDICTION, de l'EXTRAPOLATION lorsqu' elles entendent proposer des résultats concernant une population plus vaste que la population sur laquelle on dispose de données (ce que l'on nomme l'échantillon).

- elles sont de l'ordre de la CREATION d'informations lorsqu'elles révèlent une corrélation entre deux variables.

.....I..... LES STATISTIQUES DESCRIPTIVES.

...1. La collecte des données.

Contrairement à une idée parfois répandue, l'étude statistique d'un phénomène ne commence pas par collecter autant de chiffres que l'on peut en remettant à plus tard la question de l'usage qu'on pourra en faire.

Il faut , en préalable à toute collecte, se poser la question : quelles sont les données pertinentes au problème que l'on se pose?

Pour le consommateur de statistiques, il serait intéressant d'avoir un aperçu des données que le statisticien a jugé inappropriées afin d'entrer dans le sujet de manière active:

Ex: nous n'avons pas pris en compte le taux de réussite au bac parce que nous voulions étudier la réussite au lycée et non la réussite en terminale.

Malheureusement on trouve rarement ce genre d'entrées en matière dans les articles rédigés par les médias, les seuls que nous consommons régulièrement.

2. Comment décrire des données numériques concernant un seul paramètre.

- tout d'abord en les ordonnant (du plus petit au plus grand par exemple) Nous évitons ici d'utiliser le verbe "classer" qui est ambigü. Il signifie à la fois ordonner (on classe des élèves), mais aussi faire des regroupements (des classes...) .(Quand on classe ses papiers dans des chemises, on ne les range pas du plus petit au plus grand...)

- ensuite en faisant des groupements pour RESUMER les données,pour les réduire. (Mais évidemment, le choix des seuils est souvent artificiel et introduit des discontinuités parasites. Ainsi une population est urbaine si elle habite une ville de plus de 2000 habitants... Une ville pourra passer de 1900 à 2050 habitants sans que son paysage soit modifié et pourtant elle aura changé de classe...) Là aussi, un consommateur pourrait être informé du pourquoi des choix des seuils...Sont-ils placés à intervalles réguliers ou bien sont-ils adaptés à des regroupements constatés des données?

Ex, si dans une entreprise , il se trouve que le tiers des employés a entre 37 et 43ans, serait-il judicieux de faire un partage en classes de 20 à 30 ans, de 31 à 40, de 41 à 50... fractionnant artificiellement le groupement réel constaté?

- Si on étudie la consommation d'énergie par habitant (en kg d'équivalent pétrole),on obtient des chiffres comme:

8 , 9 , 11 , 15 , 95 , 98 , 100 , 110 , 3500 , 3800 , 4000 ,...

alors l'ordre de grandeur des nombre prend le pas sur leur rangement en ordre croissant. On constituera donc des classes de pays ayant des consommations voisines, puis on ordonnera les classes.

3. Réduction de données numériques à un seul chiffre:

a. la MOYENNE.

Ex: on ajoute les tailles de tous les individus de la population étudiée, et on divise par leur nombre pour obtenir la taille moyenne...

Cet indicateur tient compte de toutes les mesures mais il est influencé par les valeurs extrêmes. Il n'a pas beaucoup de sens si les valeurs sont très dispersées. (Si les notes d'un élève sont dans la fourchette 8 - 12 et si sa moyenne est 10,5, ce nombre reflète mieux la réalité que s'il s'agit d'un élève ayant essentiellement des 5 et des 16...)

Se demander aussi par exemple quel est le sens de "la taille moyenne des français"..lesquels sont constitués de deux populations disjointes (hommes- femmes) ayant une spécificité différente à ce sujet...

b. la MEDIANE.

Si on dit que la moitié des salariés d'une entreprise gagne moins de 5000f par mois, on se réfère à une valeur appelée la médiane (celle qui partage l'effectif en deux moitiés égales). La médiane n'a aucune raison d'être égale à la moyenne. (Ainsi quelques trés gros salaires feront monter la moyenne...) .L'intérêt de la

médiane est précisément de ne pas être influencée par les valeurs extrêmes. Mais d'une certaine façon on peut dire aussi qu'elle ne prend pas en compte toutes les valeurs... Deux entreprises peuvent avoir la même médiane et des réalités bien différentes (si la moitié des salariés gagnent 4950 f...la médiane donne une vue pessimiste de la réalité, laquelle n'est cependant pas réjouissante d'ailleurs!)

c. le MODE.

C'est la valeur la plus souvent atteinte par le paramètre étudié. On pourrait dire que c'est l'indicateur des commerçants! En effet si un marchand de chaussures de femmes constate que la pointure la plus vendue est le 37, il réapprovisionnera ses stocks en conséquence On voit mal l'usage qu'il pourrait faire de la pointure moyenne...!

Pour la vente des tennis, on aurait une répartition bimodale avec deux pointes, un mode pour les femmes et un mode pour les hommes.

On voit sur ces exemples simples qu'un seul chiffre pour résumer des données nombreuses, risque toujours d'induire des erreurs. Connaître à la fois la moyenne et la médiane permet déjà de mieux cerner la réalité.

4. Les indicateurs de dispersion.

- la fourchette des valeurs extrêmes est l'indicateur le plus simple concernant la dispersion des données. C'est ce que l'on fait quand on indique la note la plus haute et la plus basse de la classe...sauf que là on donne la dispersion des notes données par l'enseignant et non celle des notes obtenues par l'élève! En fait les deux seraient utiles! (en particulier lorsqu'une décision est prise au seul vu des notes, en commission d'appel. Moyenne et fourchette des notes du prof dans l'ensemble de la classe, puis moyenne et fourchette des notes de l'élève seraient complémentaires.)

- L'indicateur le plus utilisé pour la dispersion est l'écart type.

Lorsque la population étudiée est nombreuse et que le paramètre suit une loi "en cloche" (les valeurs extrêmes sont peu souvent atteintes, c'est le cas pour la taille des françaises par exemple), alors l'écart type a une interprétation concrète bien précise.

Supposons que la taille moyenne des françaises soit de 163cm avec un écart type de 7 cm, et qu'en outre cette variable suive une loi "normale" (ou loi de Gauss, ou courbe en cloche, tout cela veut dire la même chose), alors cela veut dire:

- que 68% des françaises mesurent entre 163 - 7 et 163+7, c'est-à-dire entre 156 et 170 cm,

- et que 95% des françaises mesurent entre 163 - 2.7 et 163 + 2.7, soit entre 149 et 177 cm.

Pour toutes les lois en cloche régulière, il y a toujours 68% de l'effectif dans la fourchette:

(moyenne - 1 écart-type) - - - (moyenne + 1 é type)

et 95 % dans la fourchette :

(moyenne - 2 écart-type) - - - (moyenne + 2 écart-type)

Exemple: Voici comment on est "statistiquement malade"!

Institut Pasteur de Lyon

EXPLORATION THYROIDIENNE

TEST DE FIXATION DE T3 .....: supérieur à 0,60
Valeurs normales: 0,34-0,46 (m ± 2 DS)

T4 (nmol/l).....: 249
Méthode radioimmunologique
Valeurs normales: 70-130 nmol/l (m ± 2 DS)

Le laboratoire indique entre parenthèses que la fourchette normale correspond à une fourchette de type moyenne plus ou moins 2 ecarts - type . C'est cela qui donne son sens au chiffre obtenu par le patient. Si le chiffre est hors de la fourchette, le patient n'est pas dans les 95% de "normaux". Il peut être soit dans les 2,5% qui fonctionnent trop (c'est le cas ici), soit dans les 2,5% qui ne fonctionnent pas assez. Dans les deux cas il sera considéré comme...à soigner!

Il existe une table numérique (en une page) permettant de dire, au vu du résultat obtenu par le patient, s'il est plus précisément dans les 1% ou 0,5%,ou 0,1% qui fonctionne trop. C'est-à-dire qu'on peut estimer le degré de rareté de son cas.

Cette table, dite table de la loi normale centrée réduite est universelle et peut concerner des situations totalement différentes dans leur contexte:

Ex : si une entreprise de magnétoscopes a établi,par expérience passée, que la durée de fonctionnement sans panne de ses appareils est de 2ans avec un écart type de 5 mois, cela veut dire que 95% des appareils auront leur première panne entre 14 mois et 34 mois. 2,5 % auront

leur première panne après 34 mois...et 2,5% auront leur première panne avant 14 mois. Donc en tenant compte du prix moyen des premières pannes, on peut calculer la dépense maximum envisageable pour les 2,5 % et répartir ces frais sur la totalité des appareils vendus. En mettant une garantie d'un an, l'entreprise est sûre que les frais de réparation sous-garantie seront couverts par le prix de vente...

Nous n'en dirons pas plus ici. Le livre " Statistiques, Une approche nouvelle " donne de nombreux exemples d'utilisation de la table de la loi normale.

Il donne aussi la formule pour calculer ce fameux écart-type, formule connue de tous les candidats bacheliers... formule que j'ai enseignée plusieurs années sans savoir ce que représentait l'écart-type... Cela me semble un très bon exemple de disparition de sens d'une notion lorsqu'elle pénétre dans le "scolaire"... en particulier quand elle a l'insigne honneur de figurer dans un programme de mathématiques! Donner du sens, pourquoi faire ?

..... II ..... STATISTIQUES D'EXTRAPOLATION.

Prenons un exemple caricatural pour comprendre.

Si on fabrique un échantillon de 1000 personnes grenobloises en respectant

- les proportions concernant les âges,
- les proportions concernant les catégories socio-professionnelles,
- les proportions concernant les religions,
- etc...

Et si on interroge cet échantillon, modèle réduit de la population grenobloise, sur sa consommation de tabac, on pourra savoir avec certitude que, par exemple il y a 30% de fumeurs dans l'échantillon étudié.

On ne pourra jamais dire AVEC CERTITUDE qu'il y a 30% de fumeurs dans l'ensemble de la population grenobloise.

On pourra dire avec une certaine certitude (qu'on appelle un degré de confiance) qu'il y a entre 28 et 32% de fumeurs dans la population totale,

On pourra dire avec une certitude plus grande que la précédente qu'il y a entre 25 et 35% de fumeurs dans la population totale...

et avec une certitude encore plus grande qu'il y a entre 20 et 40% de fumeurs..

(et avec une certitude totale qu'il y en a entre 0% et 100% !)

Tout ce que je souhaite faire comprendre ici, c'est le jeu de balance entre le niveau de confiance et l'amplitude de la fourchette.

Le niveau de confiance est gradué de O à 100%. Avoir un degré de confiance de o,8 veut dire: j'ai 80 chances sur 100 que mon affirmation soit juste (et évidemment, 20% de chances qu'elle soit fausse...)

Trop élargir la fourchette des résultats pour avoir un haut niveau de confiance aboutit à énoncer...une évidence, et donc à avoir gaspillé temps et argent à un sondage inutile!

Trop restreindre la fourchette pour avoir un résultat "précis" fait augmenter le risque d'erreur... Il y a donc des choix à faire.

Ajoutons pour être complet que la taille de l'échantillon entre aussi en ligne de compte. La relation qui lie le degré de confiance et la largeur de la fourchette n'est pas la même pour des échantillons de 1000 ou 3000 individus... Plus la taille de l'échantillon est grande, plus le degré de confiance est grand pour une même fourchette... A la limite, si on interroge TOUTE la population (donc le plus grand échantillon possible!), alors on atteint la certitude! Mais ce ne sont plus des statistiques d'extrapolation, ce n'est plus un sondage mais une enquête.

(En principe on utilise le mot enquête quand on interroge <u>toute</u> la population dont on veut parler, et le mot sondage si on interroge seulement un échantillon).

Depuis plusieurs années, les sondages qui paraissent dans les journaux nous indiquent la taille de l'échantillon,mais à vrai dire le consommateur ne sait trop quoi faire de ce renseignement (cela évite tout de même des abus : des échantillons de 30 personnes! Mais attention, on vous donne toujours la taille totale de l'échantillon, et souvent les questions proposent un découpage de l'échantillon: par exemple : les moins de 25 ans, ou les femmes de moins de 30 ans déjà mère de famille etc... On risque alors, dans ce sous-découpage, de rencontrer des échantillons de tailles trop réduites...et le degré de certitude tombe alors en chute libre!)

Il vaudrait mieux que les résultats soient francs , du genre:

Nous pouvons affirmer (avec 10% de chances de nous tromper) qu'il y a dans Grenoble entre 28 et 32% de fumeurs....

Très rares sont les journaux qui osent cette forme de franchise, ou qui osent une réflexion sur les résultats statistiques. On pourra lire à la fin de ce chapitre un article intéressant à ce propos.

Le choix du degré de confiance voulu, de la taille de l'échantillon, de l'amplitude de la fourchette dépend du contexte...et du budget disponible! IL est bien évident que si des intentions de vote tournent autour de 52% dans une élection comportant 2 candidats, il ne faudra pas se baser sur une fourchette du genre 48% - 56% qui enlève toute interprétation possible au résultat que l'on cherche!

Voilà tout ce que nous dirons à propos du passage d'un échantillon à une population totale. (Sur ce sujet, le livre canadien déja cité est riche, mais c'est un peu plus difficile à comprendre. Heureusement qu'il y a les exercices avec solution! Il y a cependant un défaut dans le plan proposé pour cette étude, car on se demande pendant 40 pages comment un chercheur peut utiliser une estimation du résultat qu'il cherche avant de l'avoir trouvé... et les exercices proposés à ce sujet paraissent dénués de sens: on a l'impression de chercher ce que l'on sait par hypothèse...Il s'agit bien sûr d'une impression personnelle.)

Mais n'oublions pas non plus que le passage d'une population totale à un échantillon pose le même genre de questions, bien que cela paraisse à première vue paradoxal.

En effet, si l'on sait après enquête exhaustive d'une population de donneurs de sang que 98% des donneurs ne sont pas contaminés par le virus du SIDA, il n'empêche que si l'on prenait un échantillon de 100 donneurs, au hasard, dans cette population, on pourrait trouver 90% de donneurs sains ou bien 100%...

Chaque échantillon de 100 personnes aurait son taux de donneurs sains, et avec tous les échantillons différents que l'on pourrait faire, on aurait une distribution de valeurs des taux réels correspondants qui suivrait une loi "en cloche", et dont la moyenne serait précisément de 98%...

De ce fait, prenant un échantillon de 100 personnes, au hasard, on pourra dire aussi avec un degré de confiance connu dans quelle fourchette est situé l'échantillon. On ne peut réellement donner qu'un résultat sous la forme:

Pour cet échantillon, nous sommes sûr (mettons à 95%) que le taux de donneurs sains sera compris entre 97 et 99 %.

..... III ..... CORRELATION STATISTIQUE.

Etudier la corrélation entre deux événements, c'est se demander s'ils se produisent souvent -voire très souvent - ensemble, ou rarement - voire très rarement - ensemble, ou bien encore si l'apparition de l'un est indépendante de l'apparition de l'autre.

On dira qu'il existe une forte corrélation dans les deux premiers cas , corrélation positive qui rejoint la notion d'attraction vue en analyse factorielle(*) et corrélation négative pour la répulsion. Mais dans les deux cas de forte corrélation, on pourra en déduire une prédiction au sens propabiliste.

Ainsi, des études statistiques ont montré une assez forte corrélation positive entre l'usage du tabac chez la mère et la prématurité de l'enfant à naître. Cela veut dire qu'une future mère grande fumeuse peut s'attendre, plus qu'une autre, à ne pas mener sa grossesse à terme. Mais cette prédiction reste du domaine des probabilités En langage courant, on dira qu'elle a moins de chance (que si elle ne fumait pas) d'arriver au terme des neuf mois.

Ces résultats basés sur la statistique laissent toujours sceptiques car chacun connaît des exceptions. Mais

- d'une part, cela n'est pas contradictoire avec le résultat statistique qui ne nie pas l'exception. Au contraire il la reconnaît en lui attribuant un faible pourcentage de chances de se réaliser.

- d'autre part il rectifie notre vision qui n'est pas neutre : on se souvient mieux des exceptions parce qu'elles semblent paradoxales.

> *" En effet, en l'absence de tels dénombrements, on risque beaucoup de ne raisonner que sur des cas exceptionnels et de négliger les cas les plus fréquents.(...) C'est ainsi que l'on serait conduit à dire que les échecs scolaires sont associés à une brillante réussite professionnelle ultérieure; que les émigrants arrivant sans ressource aux Etats-Unis y font des fortunes colossales.."*

Cette citation est extraite du Précis de Statistique, de Maurice Reuchlin, édité aux PUF,collection Le Psychologue,page 112 .

* Pour l'analyse factorielle, voir page 67.

Nous avons une mémoire qui privilégie l'exception, ce qui est somme toute normal, l'exception est plus "frappante", il faut le savoir. Accepter un résultat statistique ,c'est augmenter sa lucidité et ne pas se forger une opinion à partir d'une vision déformée de la réalité.

Un exemple vécu, très caractéristique: Une enseignante de mathématiques Spéciales affirmant par écrit dans la salle des professeurs, que l'admission en classe préparatoire ne dépendait pas du tout de la profession des parents, la preuve , un fils de cadre supérieur avait été refusé et un fils d'ouvrier accepté. (Je ne caricature pas! et il s'agissait d'une "prof. de math.") Elle refusa le résultat statistique... "de toutes façons, les nombres disent n'importe quoi." Je me souviens encore de ma "rage" à voir une collègue se comporter aussi peu scientifiquement. Et je sentais bien que, lors de notre discussion, c'est son discours qui emportait l'adhésion des collègues présents...

L'éducation au langage numérique est un long marathon,je l'admets mieux aujourd'hui...

Il faut aussi apprendre à se méfier des fausses apparences de vérité à propos de résultats dichotomiques pour lesquels chaque éventualité a une chance sur deux de se produire. C'est l'exemple du pile ou face de la pièce de monnaie, mais aussi des naissances de garçons et de filles.

On peut très bien faire croire qu'on devine le sexe d'un enfant à naître en s'inventant une règle sur le principe "une chance sur deux". Exemple : si la mère est née dans les 15 premiers jours du mois, ce sera un garçon..., si le père est né un jour pair ce sera une fille..., si la mère est née un mois pair ce sera un garçon... etc. Dans tous ces cas (et sur un nombre assez grand d'exemples),on prédira juste pour 50% des prédictions...Cela n'est certes pas terrible, mais on pourra au moins citer de "nombreux témoignages" - d'ailleurs réels - de réussite et donc recruter des clients naïfs et crédules...

*** Mesure de la corrélation.

Les statisticiens ont inventé une mesure de la corrélation, un coefficient variant de -1 à +1. Mais on peut reconnaître à l 'oeil nu s'il y a corrélation positive, négative, ou indépendance des deux variables étudiées.

Cela consiste à construire les nuages de points.

Prenons un exemple facile à réaliser par des enseignants.

Vous demandez à vos élèves leur note de français et de géographie. Et pour chaque élève vous placez un point. Si vous faites cela pour plusieurs classes (pour ne pas limiter l'échantillon à 35), vous obtenez un nuage:

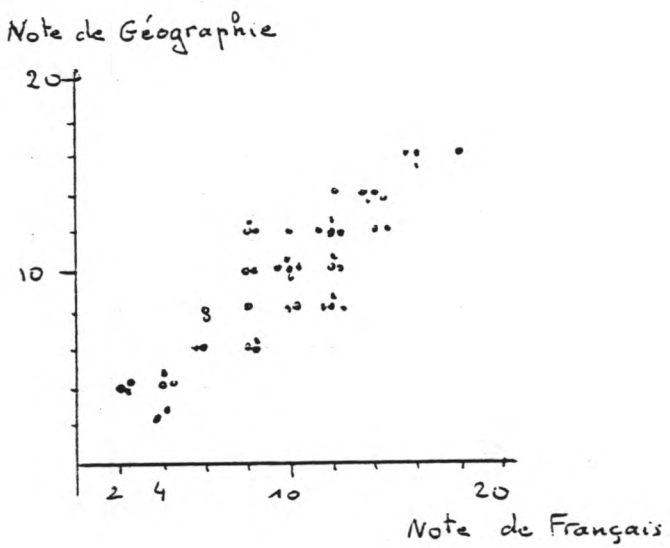


figure 1

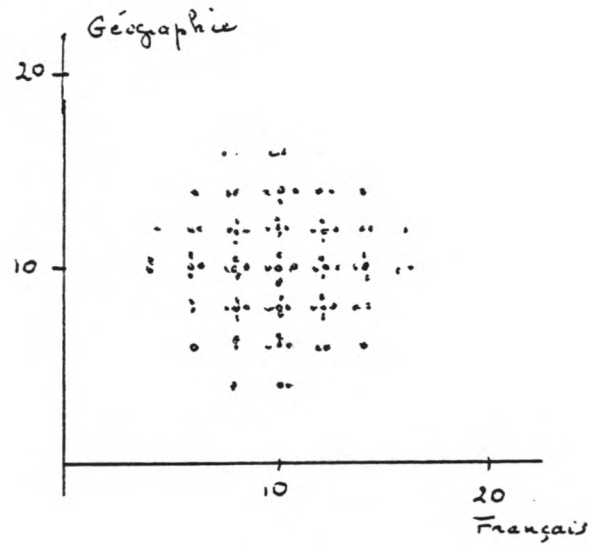


figure 2

Si le nuage est étiré comme dans la figure 1, il y a une assez forte corrélation. (Ici, une corrélation positive : quand l'une des notes monte, l'autre aussi)

Si au contraire le nuage est en rond, il y a indépendance, c'est à dire corrélation nulle. (figure 2)

Bien sûr il y a tous les intermédiaires... Le coefficient de corrélation mesure justement le degré d'étirement rectiligne du nuage (car , nous ne l'étudierons pas ici, mais des nuages peuvent être étirés en forme de "banane". Voir le chapitre Graphiques logarithmiques dans le fascicule Graphiques.)

La corrélation négative se traduirait par un nuage semblable à celui de la figure 3.

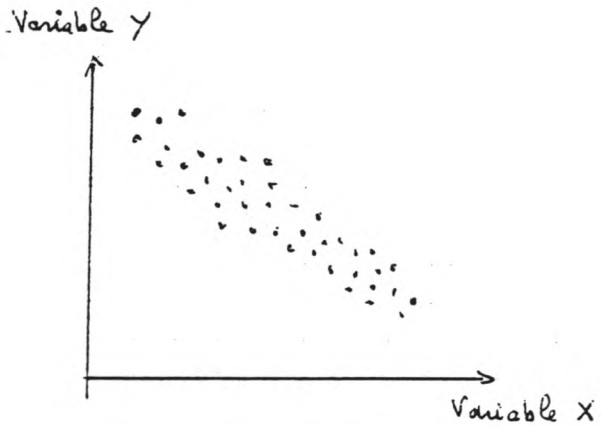
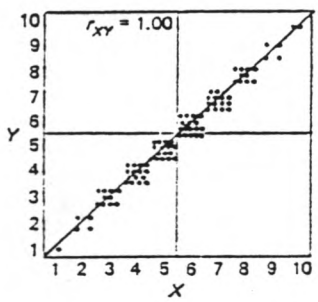
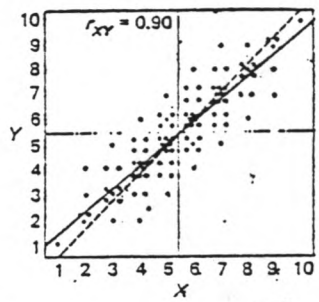


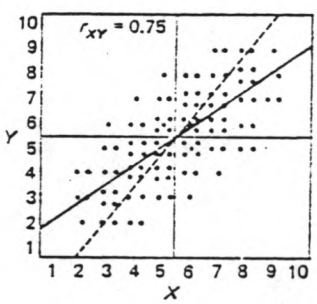
figure 3.

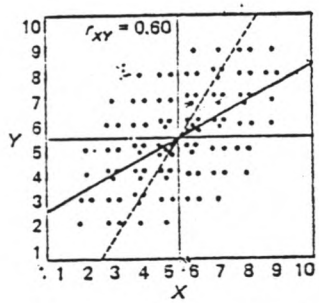
J'ai vu des nuages réalisés avec la note de math de troisième et la note de math de seconde (par l'IREM de Besançon,je crois, mais impossible de remettre la main dessus! Eternel problème des documents susceptibles d'intéresser plusieurs rubriques...) Cela dit,les nuages en questions ne montraient pas une forte corrélation!

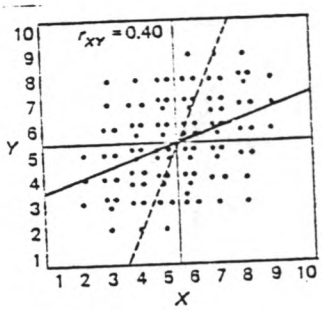
Voici des exemples de nuages et de valeurs du coefficient de correlation correspondant , noté r_{xy}

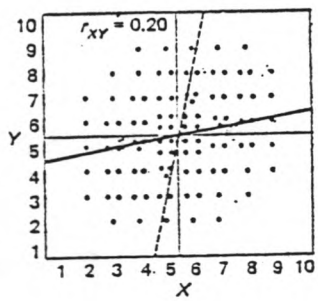


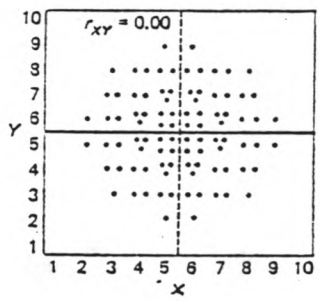










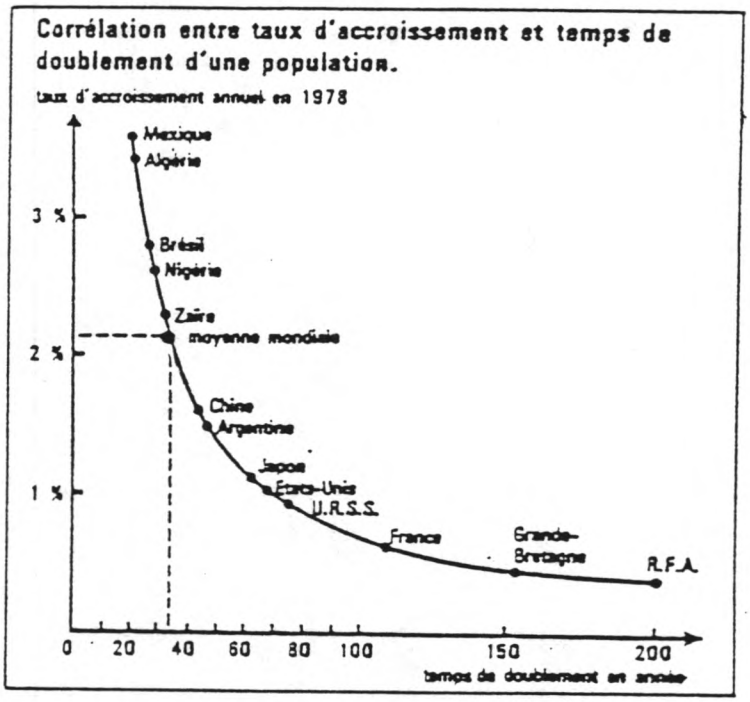


précis de statistique

maurice reuchlin

Quand le nuage est bien allongé suivant une droite,on peut le "résumer" précisément, par une droite, appelée la droite de régression. Mais lorsque le nuage est moins étiré,cela devient de moins en moins possible... la droite se dédouble en deux droites de régression suivant que l'on considère que x dépend de Y, ou y dépend de x. Ces deux droites doivent résumer la direction globale du nuage. On voit que pour un nuage rond (coefficient nul) ,cela n'a plus de sens de dessiner ces deux droites puisqu'elles sont orthogonales! Elles n'indiquent plus de direction globale du nuage, ce qui est normal puisqu'étant rond, il n'en a pas.

*** La corrélation n'est pas une liaison mathématique.



Ce temps est d'autant plus faible que le taux est élevé : 25 ans pour un taux de 3 % et près de 200 ans pour les faibles taux d'Europe (lorsqu'ils sont positifs). L'ensemble des valeurs obtenues se classe selon une fonction hyperbolique, les deux paramètres variant en fonction inverse.

Voici un bel exemple de confusion, trouvé dans un manuel de géographie.

Manuel de géographie
Editions Bordas,
Classe de seconde.
(page 21)

Le commentaire parle d'ailleurs très justement de FONCTION. (Dans le fascicule Graphiques, nous montrons comment on peut faire retrouver la fonction à des élèves de seconde.)

Puisqu'il y a un lien fonctionnel entre le taux d'accroissement t et le temps de doublement n (n années), il y a certitude. Moyennant une approximation, le lien se résume en t.n = 1 , donc les points correspondants à chaque pays se disposent exactement sur une courbe. Il ne s'agit pas d'un nuage de points. A la limite, on pourrait dire que la corrélation est maximum (100%), mais je trouve préférable de ne pas employer ce terme quand il y a un lien de fonction. Gardons ce terme pour un lien statistique constaté expérimentalement.

*** Cause et corrélation.

Nous terminerons cette partie concernant les statistiques, par la plus grave confusion rencontrée. Une forte corrélation exprime peut-être un lien de cause à effet, mais elle ne le prouve jamais.

Dans son livre "Au péril de la science" (Seuil. Collection Points Sciences), Albert Jacquard donne un bel exemple de corrélation qui n'est pas une cause.

Il nous propose , en effet, page 95, un exemple un peu caricatural:

> *" Traçons la courbe mensuelle de la consommation de charbon, puis , sur le même graphique, la courbe de mortalité des vieillards; les deux courbes sont remarquablement parallèles, elles marquent une pointe en janvier-février, une chute en juin - juillet. Faut-il en conclure que la diminution de la consommation du charbon ferait décroître le nombre de décès de vieillards?"*

Bien sûr, sur cet exemple, tout le monde reconnaît bien que les deux phénomènes ont une cause commune, le froid. Mais cela n'est pas toujours aussi simple...

Ainsi pour l'exemple de la prématurité et la consommation de tabac, certains médecins n'ont pas hésité à parler de lien de cause à effet à leur cliente, alors que les deux faits peuvent être analysés comme symptômes concomittants, signes d'une nervosité ou d'une anxiété de la mère...de cause non forcément connue. Le symptôme "tabac", plus mesurable, peut permettre de déceler des futures mères à risque et donc de leur proposer un meilleur suivi de leur grossesse.

Méfions-nous donc des déductions hâtives, surtout lorsque l'une des variables semble pouvoir "expliquer" l'autre. En effet,ce n'est sûrement pas parce que la femme a un enfant prématuré qu'elle est "grande fumeuse"...avant l'accouchement! Par contre, le décalage dans le temps peut donner envie d'affirmer l'inverse: c'est parce qu'elle a beaucoup fumé qu'elle n'a pu mener sa grossesse à terme.

Retenons donc qu'un lien de corrélation n'est pas un lien de cause à effet.

* * * *

En espèrant que ces quelques pages vous auront donné un aperçu de l'état d'esprit statistique, pour reprendre l'expression d'un stagiaire géographe de Grenoble....

Page 42 LE MONDE — Jeudi 5 décembre 1985 •••

ÉTRANGER

« Le secteur privé ne s'intéresse pas à la réalité de l'activité économique mais à ses variations »

affirme M. Sidney Jones, ancien sous-secrétaire d'Etat américain au commerce

Washington. — Le monde financier américain est suspendu, à la décimale près, à la publication des statistiques d'activité de l'économie américaine. Le caractère volatile des anticipations est telle que la moindre variation de tel ou tel indice peut entraîner un retournement du marché. On peut toutefois se demander dans quelle mesure ces statistiques sont réellement significatives et comment elles sont préparées avant d'être publiées.

Jusqu'à ce qu'il quitte ce poste, il y a quelques jours, M. Sidney Jones, sous-secrétaire d'Etat au commerce pour les affaires économiques, a été le principal responsable de la présentation de ces données. Après seize années passées dans l'administration, il a choisi de retourner à son ancien métier de professeur et de travailler à nouveau pour le secteur privé. Unanimement respecté, M. Jones a toujours fait preuve d'une grande indépendance. Il vient de redevenir libre de ses propos.

Correspondance

« Nous disposons aux Etats-Unis d'un système d'information statistiques sophistiqué. Le Bureau of Economic Analysis rassemble les données et le National Bureau of Census les traite. Tous deux font partie du ministère du commerce. En outre, des ministères comme celui du travail ou de l'agriculture ont chacun leur propre système de statistiques.

» Mais il est vrai qu'à l'heure actuelle ce système nous pose des problèmes. D'abord, il est considéré par ceux qui en ont la charge comme une terrible contrainte. Ensuite, la nature même de ces statistiques fait que lorsqu'on les combine pour prendre des décisions politiques, par exemple pour la préparation du budget, il faut reconnaître qu'on échafaude des raisonnements subtils sur des données fausses.

— Et quand la Bourse s'écroule sur l'annonce d'une fuite indiquant un taux de croissance du PNB de 1,7 % au lieu des 2,5 % attendus, comme cela vous est arrivé en juillet dernier ?

— C'est encore pis. Le secteur privé ne s'intéresse pas à la réalité de l'activité économique, mais à ses variations. Les intermédiaires ne vivent que des différences qui entraînent des opérations sur lesquelles ils ont des commissions.

— Quel est le meilleur exemple de mauvaises décisions prises sur des statistiques précises ?

— Le pire exemple est incontestablement celui des statistiques mensuelles du commerce extérieur, importations et exportations, d'où se déduit le déficit commercial. Si vous observez les importations annoncées en janvier dernier, qui sont de 30 milliards de dollars, 16 milliards sont en fait des opérations ayant eu lieu dans les mois précédents, mais dont les documents sont arrivés en retard au Bureau of Census.

— Que faire ?

— J'ai pris conscience du problème en août dernier. Nous avons repris dix-huit mois de données du commerce extérieur, ce qui nous a conduits à traiter vingt millions de documents pendant l'été. Le résultat final n'est guère différent de celui annoncé. 123 milliards de dollars de déficit au lieu de 126, mais si les renseignements avaient pu être donnés plus régulièrement, ou moins souvent, et avec soixante jours de délai par rapport aux opérations au lieu de trente jours comme nous y sommes tenus actuellement, les variations n'auraient pas été si grandes et nous auraient évité des discours inutiles.

— Quelle est l'importance réelle de cette statistique du commerce extérieur ?

— L'ennui est que c'est l'une des statistiques majeures qui entrent dans le calcul de la croissance.

» En deux mots, le PNB est calculé à partir de six éléments : la consommation des ménages, les investissements des entreprises, le logement, la variation des stocks, la variation des exportations nettes et les dépenses de l'administration. Déjà deux de ces éléments sont par définition des variations, celles du commerce extérieur et des stocks. Les deux sont aussi approximatifs pour des raisons voisines.

» Lorsque vous regardez les résultats de la croissance, que nous sommes tenus d'annoncer trimestriellement, à partir du second trimestre de 1984 par exemple, vous avez l'impression que l'économie américaine est en pleine agitation. 7,1 % de croissance au deuxième trimestre, 1,6 % au troisième, 4,6 % au quatrième, 0,3 % au premier trimestre de 1985, 1,9 % au deuxième et 4,3 % au troisième. Or il n'est pas vraisemblable que les éléments réels changent aussi vite. Là encore, c'est qu'il s'agit de la mesure des variations et non de la mesure des niveaux. En somme, il s'agit de la variation des variations. Il faut au moins un mois pour expliquer la différence à des étudiants en sciences économiques. Je n'ai pas cessé d'essayer de le faire comprendre au marché, mais personne ne veut entendre.

Révision

— Les Etats-Unis approchent d'une grande période de révision statistique. Quelles vont être ses conséquences ?

— En effet. Le 23 décembre, la grande révision générale des statistiques sera annoncée. Nous avons une révision tous les cinq ans par le Bureau of Census, et une révision tous les dix ans de l'ensemble national. C'est la plus importante. Tous les indices de référence vont changer. Les nouveaux indices vont inclure les nouveaux prix du pétrole de la fin des années 70. Toutes les mesures de l'inflation vont changer. Nous allons commencer à travailler en dollars constants de 1982 alors que nous étions jusqu'à présent sur des dollars constants de 1972. La valeur de tous les biens va être modifiée, de ce fait.

— Cette grande révision envisage aussi des améliorations dans l'enregistrement statistique ?

— Oui. Par exemple, nous allons désormais essayer d'inclure de plus en plus l'économie souterraine dans nos calculs. Nous avons fait des études d'échantillons et cherché la relation entre les revenus et les dépenses. Les Américains dépensent régulièrement 91 % de leurs revenus, épargnent 6,5 % et paient 2,5 % en frais financiers. Nous sommes arrivés à la conclusion qu'environ 20 % des revenus ne sont pas déclarés au fisc.

» Nous allons essayer aussi d'améliorer notre perception de l'évolution de la productivité, notamment dans les services.

— Ces changements vous satisferont-ils ?

— D'une manière générale, je dirai que les économistes ont survendu nos capacités statistiques, et quoi que nous fassions, nous serons au-dessous de leurs espérances. Nos statistiques sont meilleures aujourd'hui qu'il y a dix ans, elles seront meilleures dans dix ans. Mais je déplore l'incapacité dans laquelle sont tous les intéressés à cette question de bien définir ce qu'ils veulent mesurer, et de se demander si le chiffre auquel ils se réfèrent est bien la mesure de ce qu'ils recherchent. La mauvaise interprétation des statistiques est encore pire que leur approximation. Personnellement, je trouve qu'on devrait mieux décrire les niveaux d'activité et moins les variations, publier des statistiques moins souvent et indiquer des fourchettes plus que des chiffres précis.

— Y a-t-il, dans le monde politique, des moyens d'interférer dans la présentation des statistiques ou des tentations de le faire.

— Depuis que je travaille pour le gouvernement, je n'ai jamais vu les statistiques avant de les présenter aux médias, sauf quand j'étais chargé du commentaire sur la croissance. On me les montrait quarante-cinq minutes avant la conférence de presse. Les communiqués étaient déjà prêts. Aucun fonctionnaire nommé par le pouvoir politique n'a accès à la préparation de ces statistiques. C'est le seul fait de techniciens de carrière.

» La veille, les chiffres sont communiqués au président des conseillers économiques de la Maison Blanche, qui peut en informer le président, au secrétaire au Trésor et au président de la Banque de réserve fédérale. C'est une pratique contre laquelle j'ai lutté depuis 1974, sans succès. Mais il faut dire qu'ils n'ont aucune possibilité d'influer sur les chiffres puisque les communiqués de presse sont déjà prêts. Et s'il y avait la moindre fuite, cela se remarquerait immédiatement. »

Propos recueillis par
JACQUELINE GRAPIN.

IV – INITIATION A L'ANALYSE FACTORIELLE

Le travail décrit ici a été fait avec une classe de seconde en fin d'année scolaire, ce qui explique sans aucun doute certaines initiatives des élèves face aux tableaux numériques.

*** Exemple de traitement d'un tableau numérique.

Les chiffres que nous allons étudier sont extraits du livre d'Alain Leger : "Enseignants du secondaire". Cet auteur est un chercheur en Sciences de l'Education. (éditions PUF ; L'éducateur)

A. Leger veut prouver que, dès qu'ils le peuvent, c'est-à-dire dès que leur ancienneté le leur permet, les enseignants se font nommer dans des établissements scolaires de plus en plus bourgeois.

Il propose à l'appui de son hypothèse un certain nombre de chiffres dont il serait logique de penser qu'ils sont destinés à convaincre le lecteur de la vérité de l'hypothèse.

LES DONNEES :

L'étude porte sur 4553 postes répartis dans des établissements, lesquels ont été regroupés en 5 classes:

La description des postes figure en annexe, page 230.

Il y a en tout 4553 postes.

2161 (soit 47,5%) sont en établissement "très bourgeois".
1432 (soit 31,5%) sont en établissement "bourgeois"
567 (soit 12,5%) sont en établissement "socialement mixte"
245 (soit 5,4%) sont en établissement "populaire"
146 (soit 3,3%) sont en établissement "très populaire".

(Le pourcentage d'enfants d'ouvriers dans la population française âgée de 18 ans en 75 étant de 45% d'après l'INSEE, A. Léger a classé les établissements en "très bourgeois" s'ils ont moins de 30% d'enfants d'ouvriers parmi leurs élèves, "Bourgeois" si cela va de 30 à 39%, "Mixte" pour 40 à 49%, "Populaire" pour 50 à 59%, et enfin très populaire pour ceux qui ont plus de 60% d'enfants d'ouvriers.

Voici le tableau donné dans le corps du texte afin de prouver le discours. Nous ne reportons ici que les chiffres concernant les 2725 Capésiens de l'échantillon d'enseignants étudié.

Age	Trés Bourgeois	Bourgeois	Mixte	Populaire	Trés populaire	Effectif
50 ans et+	57,1	29,7	10,4	2,2	0,6	462
45-49 ans	58,8	30,5	7,7	2,6	0,4	233
40-44 ans	49,3	36,7	9,6	3	1,5	335
35 39 ans	37,6	37,8	15,7	6,1	2,8	492
30-34 ans	30,1	36,2	21,3	8	4,4	748
- de 30 ans	25,3	34,1	19,1	11,4	10,1	455

Précisons que les résultats concernant les agrégés étaient intercalés entre chacune des lignes concernant les capésiens, ce qui ne facilite pas du tout la lecture. Mais on indiquait clairement que le total des pourcentages de chaque LIGNE fait 100, ce que je n'ai pas remis, faute de place.

Comment à travers ce tableau peut-on se rendre compte de la migration des enseignants vers les lycées bourgeois, au fur et à mesure qu'ils vieillissent? Les effectifs par âge sont trés dissemblables (dans un rapport de 1 à 3) et les postes proposés le sont aussi, de 146 dans les établissements très populaires à plus de 2000 dans les trés bourgeois....

Comment comparer si "tout varie" ?

Il y eut un moment de perplexité!

Je posai alors la question autrement : Comment pourrait-on voir si les capésiens sont sur-représentés ou sous-représentés dans certaines cases ? C'est Luce, future élève de la section B,qui réagit alors : "il faudrait voir une répartition normale".

Ça serait quoi, une répartition normale ?

Luce propose alors le mode de calcul, elle est approuvée par la classe, mais nul élève ne prononce ici le mot proportionnel...

Il n'empêche que Luce a proposé l'idée fondamentale :pour pouvoir tirer une conclusion , il faut pouvoir comparer, il faut donc une référence. Ce sera donc le tableau de répartition théorique,c'est-à-dire si les enseignants se répartissaient proportionnellement à leur nombre dans chaque classe d'âge et aux postes offerts.

La classe se met donc à fabriquer le tableau théorique. Le maître fait alors dégager l'idée que, dans ce cas, on peut connaître toutes les cases du tableau en connaissant seulement les deux bords, c'est-à-dire la répartition par âge et celle des postes.

On rencontre ici la notion d'INDEPENDANCE de deux caractères C'est une notion difficile que nous retrouverons dans l'interprétation des graphiques construits par analyse factorielle.

Voici le tableau théorique: II

Age	Trés Bour.	Bour.	Mixte	Pop.	Trés pop.	répartition par AGE
50 et +	8	5,3	2,1	0,9	0,5	17
45-49	4	2,7	1	0,5	0,3	8,5
40-44	5,8	3,9	1,5	0,7	0,4	12,3
35-39	8,5	5,7	2,2	0,9	0,6	18
30-34	13	8,7	3,4	1,5	0,9	27,5
-de30	7,9	5,3	2	0,9	0,5	16,7
Répartition des POSTES	47,5	31,5	12,5	5,4	3,3	100

(Les arrondis n'ont pas été faits pour se compenser...Ainsi le total de la ligne répartition des postes donne 100,2. Rien de grave pour l'usage comparatif que l'on veut faire de ce tableau.)

On peut s'arrêter alors sur les lectures possibles lorsque l'on fixe l'un des caractères, lecture sur une ligne ou lecture sur une colonne.

Gabriel est le premier à réagir sur le fait qu'on ne pourra toujours pas comparer puisque le tableau du livre n'est pas fait comme ça... En effet il ne propose qu'une lecture par ligne puisque l'addition des pourcentages d'une même colonne n'a pas de sens...

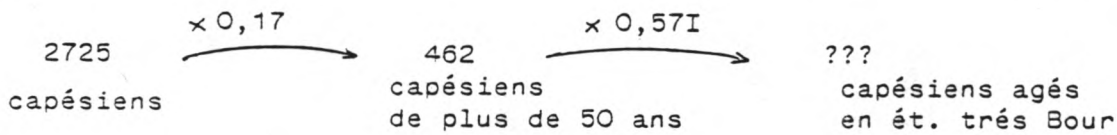
Le découragement ou l'énervement gagnent les élèves! Nous faisons une pause au soleil de juin...

(et la dactylo en fait une au soleil de décembre...)

Dans le tableau théorique, tous les pourcentages se rapportent aux 2725 enseignants,en ce qui concerne les 30 cases du tableau. Peut-on prendre le même référentiel pour fabriquer un tableau des chiffres réels?

Parmi les 462 enseignants de plus de 50 ans, 57,1% sont dans un établissement "très bourgeois". Quel pourcentage cela représente-t-il par rapport aux 2725 enseignants ?

On retrouve ici la notion de produits de pourcentages déjà développée.



0,17 x 0,571 soit o,097

(En fait la majorité des élèves préfèrent évaluer le nombre signalé par ???, puis ensuite ils le reportent sur 2725, ce qui est moins abstrait pour eux. ET comme ils peuvent se partager le travail, cela n'est pas trop long. Une minorité d'élèves verra l'intéret du calcul systématique par le produit.)

Voilà enfin un tableau des données REELLES, comparable au tableau théorique...

Age	Trés bour.	Bour.	Mixte	Pop.	Trés pop.
50 et +	9,7	5	1,8	0,4	0,1
45-49	5	2,6	0,7	0,2	0
40-44	6	4,5	1,2	0,4	0,2
35-39	6,7	6,8	2,8	1	0,5
30-34	8,3	10	5,9	2,2	1,2
-de 30	4,2	5,7	3,2	1,9	1,7

III

On pourrait alors calculer un tableau des écarts, mais pour nous reposer de tous ces chiffres, nous nous contenterons de signaler par un + (ou un ++) la sur-représentation dans une case donnée et par - (ou --) la sous-représentation.

Tableau du signe des écarts entre la répartition constatée par l'enquète et la répartition qui respecterait la proportionnalité, donc qui ne manifesterait ni "attirance" ni "répugnance":

Age	Trés Bour.	Bour.	Mixte	Pop.	Trés pop.
5o et +	+	-	-	-	-
45-49	+	O	-	-	-
40-44	O	+	-	-	-
35-39	-	+	O	o	O
30-34	-	+	+	+	O
-de 30	-	+	+	+	+

Voila un tableau qui illustre bien la migration des enseignants quand leur ancienneté le leur permet... Donc l'hypothèse d'Alain Léger est prouvée en ce qui concerne les Capésiens...

Puisque de toutes façons le tableau que nous donne l'auteur n'est pas interprétable directement...à quoi bon le donner? N'y a-t-il pas là un certain manque de respect du lecteur?

Il me semble que l'on peut très bien donner le tableau réel III, expliquer d'une phrase que l'on va chercher s'il y a ou non sur-représentation dans certaines cases, en comparant à un tableau construit par proportionnalité, puis donner le tableau des signes des écarts qui est très visuel...

Il semble que le tableau donné avec le texte ne soit là que comme "marteau" pour nous assèner l'hypothèse. D'ailleurs, pour lire réellement ce tableau, ne faut-il pas aller cueillir des données supplémentaires dans les annexes?

Nous avons mis deux heures en classe entière pour faire ce travail. L'intérêt des élèves m'a surpris, et leur indignation en fin de séquence n'était pas feinte!

Pour aborder la séquence suivante, j'ai signalé qu'il fallait retenir deux idées principales:

- Pour certains tableaux, on peut connaître entièrement les "cases" en connaissant seulement les "bords", c'est-à-dire les répartitions marginales.

- Finalement, on a inventé une opération sur des tableaux: la différence case par case. On pourrait faire de même la somme case par case. (C'est comme pour les composantes des vecteurs, a dit une voix.... Mais cela nous aurait emmené trop loin, et pourtant...)

*** Initiation à l'analyse factorielle.

Nous conseillons ici la lecture du "Que sais-je?" n° 2095 de Philippe Cibois, dont est extrait le premier exemple que nous allons regarder.

J'ai gardé le thème de cet exemple car il s'intègre bien dans les préoccupations des élèves...plus directement que les demandes de mutations des enseignants!

Donc, après "Où vont les profs ?", voici "Où vont les élèves après le bac ?".

Le tableau ci-dessous, reproduit dans le Que sais-je, émane du Service des Etudes Informatiques et Statistiques du ministère de l'Education NAtionale.(Note 76-35 du 15 octobre 76).

ORIENTATION DES BACHELIERS 75

	Université	Classes préparatoires	Autres	Total
Bac A	13	2	5	20
Bac BDD'	20	2	8	30
Bac CE	10	5	5	20
Bac Technique	7	1	22	30
Total	50	10	40	100

Comme dans l'exemple précédent, on peut faire apparaître le tableau théorique, comme si les bacheliers se répartissaient proportionnellement à leur effectif dans chaque série de bac, puis le tableau des écarts à cette situation théorique. On obtient :

	Université	Classes prépa	Autres
Bac A	+3	0	-3
Bac BDD'	+5	-1	-4
Bac CE	0	+3	-3
Tecnique	-8	-2	+10

Ce travail préliminaire peut servir d'évaluation du travail précédent, les calculs étant plus simples.

Mais l'objectif de cette séquence est d'aller plus loin, c'est à dire jusqu'à la représentation graphique des notions d'attraction, de répulsion et d'indépendance, tout en réutilisant des connaissances du programme de seconde qui ont paru sans intérêt aux élèves ne désirant pas poursuivre dans une section scientifique : les vecteurs et en particulier le produit scalaire.

Que les lecteurs non matheux se rassurent, le mode d'emploi d'un graphique d'analyse factorielle ne nécessite aucune connaissance préalable sur le produit scalaire, mais il est intéressant que les élèves, eux, se rendent compte qu'ils utilisent certaines parties du cours de math dans un domaine qui paraît pourtant fort éloigné.

Ce tableau des écarts résume l'information, mais ce n'est pas un tableau que l'on peut obtenir par la connaissance de "bords" bien choisis...

Or ce sont les tableaux réductibles aux croisements de "bords" bien trouvés , les tableaux "faciles à lire" comme les dénomme P. Cibois, qui seuls pourront être représentés sur un axe.

Ainsi le tableau

+2	-4	+6
+1	-2	+3
-3	+6	-9
2	-4	+6

(totalement inventé)

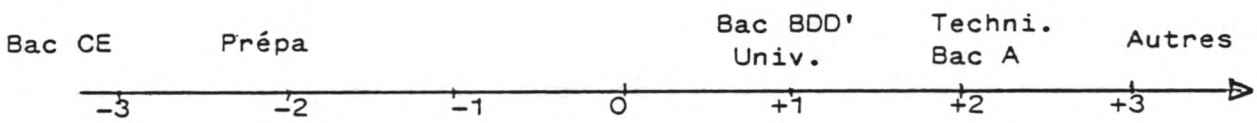
est-il en fait le produit de ses deux "bords" :

+2
+1
-3
+2

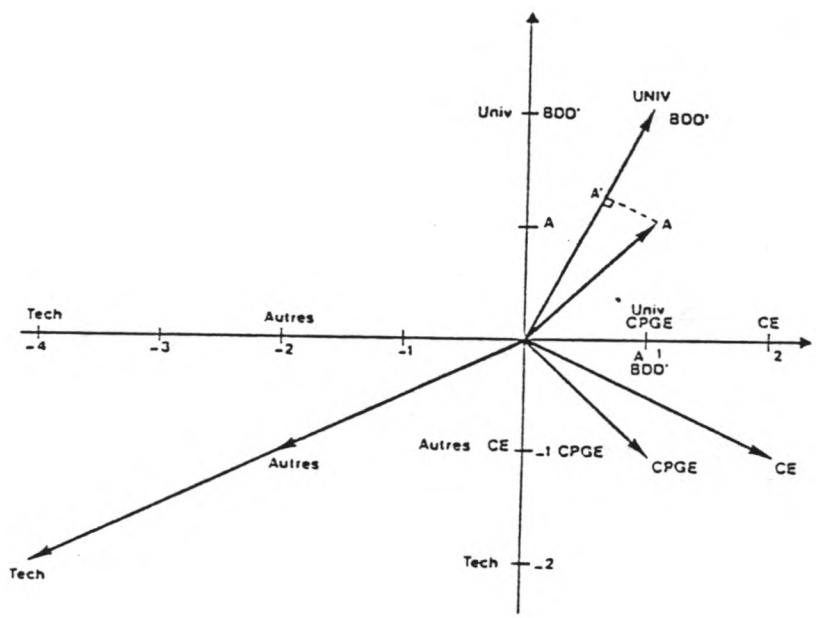
+1	-2	+3

Alors, pour représenter un tel tableau, on choisit un axe (une droite graduée) et on reporte tous les chiffres des deux bords sur le même axe, ce qui paraît choquant !

Pour ce tableau inventé, on obtiendrait:



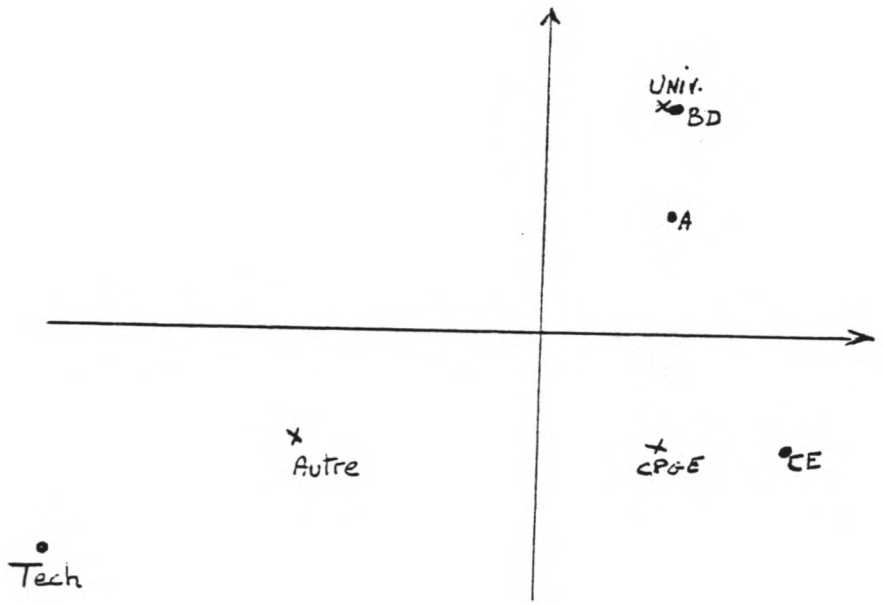
Revenons à l'exemple réel concernant l'orientation des bacheliers et au tableau des écarts obtenu. Ce tableau n'ayant que trois colonnes, on peut le ramener à une somme de deux Tableaux, chacun étant issu du croisement de ses bords. Chacun de ces tableaux peut donc être porté sur un axe. Il faudra donc deux axes puisqu'il y a deux tableaux. On les dessine orthogonaux l'un à l'autre. Sur chaque axe, il y aura donc une graduation "Université" (ici +1 sur l'axe horizontal,et +2 sur l'axe vertical), ces deux graduations seront les coordonnées d'un POINT nommé Université. Cela donne finalement:



Les lecteurs curieux de connaître les étapes intermédiaires, par exemple les tableaux dont la somme redonne le tableau des écarts, se reporteront utilement au livre de P. Cibois.

Sur le graphique précédent, des flèches sont dessinées. Or dans les manuels ou dans les journaux, on voit plutôt une représentation en points de couleurs différentes...et deux axes dont on se demande ce qu'ils viennent faire là au milieu.

On aurait par exemple, en notant avec des points les diverses séries de bac, et avec des croix les débouchés aprés le bac:



COMMENT LIRE UN TEL TABLEAU ?

Le point O, intersection des deux axes a une grande importance, car il faudra imaginer (ou dessiner) les flèches issues de O et joignant tous les points du graphique.

On pourra alors comparer une flèche de O à un point

à une flèche de O à une croix.

- Si les deux flèches tirent en sens contraire (ou presque) comme $\overrightarrow{O\ A}$ et $\overrightarrow{O\ Autres}$, cela traduit une répulsion: les élèves ayant le bac A sont sous représentés dans ce qui a été regroupé sous le label Autres.

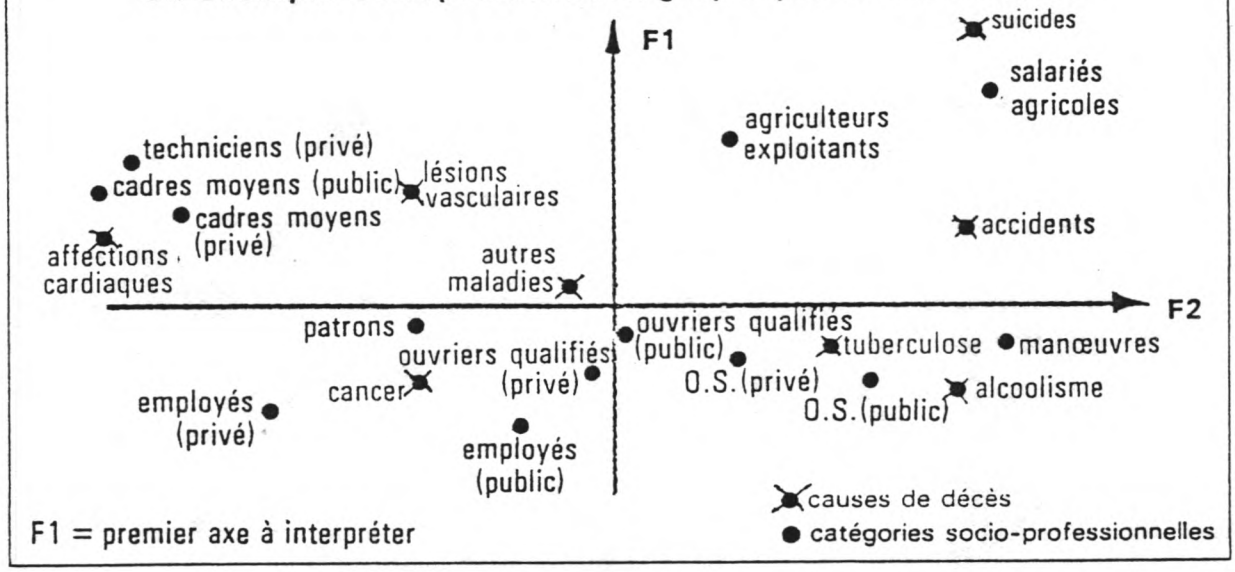
- Si les flèches tirent dans le même sens, comme $\overrightarrow{O\ BD}$ et $\overrightarrow{O\ Univ}$, cela traduit l'attraction des jeunes ayant un bac B ou D pour l'université. Ils y sont sur-représentés.

- enfin si les flèches forment un angle droit (ou presque) comme $\overrightarrow{O\ CE}$ et $\overrightarrow{O\ Univ}$, c'est qu'il y a indépendance. A l'université il y a des bacheliers CE, dans la même proportion qu'il y a de reçus au bac CE dans l'ensemble des reçus au bac.

Les enseignants de math reconnaîtront ici l'usage du produit scalaire. En fait les divers produits scalaires redonnent les divers écarts trouvés dans le tableau des écarts. En faisant la construction pas à pas, les futurs premières S ont pu retrouver la formule xx'+yy' (produit ligne.colonne ET somme de deux tableaux).

Il se trouvait que justement il y avait dans le livre de géographie de la classe un graphique de cette sorte (Géographie. Seconde chez Hatier). Laurent y pensa de suite et nous avons pu apprécier la qualité du commentaire en ce qui concerne la lecture d'un tel graphique. Les voici, extraits de la page 126:

Un exemple de représentation graphique de statistiques



Le graphique ci-contre permet d'apprécier la relation entre causes de décès et catégories socio-professionnelles en France. Il visualise, après des opérations mathématiques complexes, une information se rapportant aux 8 principales causes de décès des hommes âgés de 46 à 54 ans entre 1956 et 1960 et répartis en 13 catégories socio-professionnelles. Ces statistiques tiendraient dans un tableau à double entrée de 8 lignes et 13 colonnes, mais les relations existantes ne seraient pas visibles. Plus la distance sur le graphique entre cause de décès et catégorie socio-professionnelle est courte, plus la corrélation est grande. Ainsi, l'alcoolisme est une cause de mortalité importante chez les manœuvres et les ouvriers spécialisés, le cancer chez les patrons...

*** Prolongements hors classe...

Nous avons étudié un exemple où le nombre de lignes et de colonnes du tableau des écarts est faible, c'est un tableau 3 par 4.

Dans le cas où il y a un plus grand nombre de lignes et de colonnes, le tableau des écarts ne peut pas être ramené purement et simplement à la somme de DEUX tableaux faciles à lire...

Tableau des écarts = Tableau réductible à ses bords + tableau reste

T_E T1

On peut recommencer pour le tableau reste:

Tableau reste = Tableau réductible à ses bords + nouveau tableau reste

T2

On peut ainsi de suite ramener le tableau des écarts à:

Tableau Ecarts = T1 + T2 + T3 + etc

Nous avons dit que le tableau des écarts contient l'information à illustrer. Donc on fractionne cette information lorsqu'on décompose le tableau des écarts en somme de plusieurs tableaux , mais, et c'est très important, les tableaux successifs obtenus n'ont pas chacun une part égale de l'information contenue dans le tableau des écarts, au contraire.... C'est le premier tableau T1 qui en contient le plus, puis T2 en contient moins que T1 mais plus que T3,etc

Pour prendre une image parlante, imaginons que l'information contenue dans le tableau des écarts soit de même nature que du gravier associant des "petits cailloux" de plus en plus fin, jusqu'à du sable... T1 serait comme un tamis dont les mailles laisseraient certes échapper de l'information, mais qui ramasserait cependant "le plus gros" de la matière...On reprendrait ce qui a échappé avec un tamis à mailles plus fines, ce serait T2.

Peut-être qu'avec T1 et T2 on aurait tout ramassé...(c'est le cas de l'exemple étudié. Alors le graphique contient toute l'information).

Peut-être qu'on a encore laissé échapper de l'information, on reprend un tamis encore plus fin, ce sera T3 etc

Si l'on estime qu'avec T1 et T2 on a déjà les trois quarts de l'information, on peut s'en contenter. Alors le graphique à deux dimensions est une APPROXIMATION de l'information.

Si T1 et T2 ne nous semble pas suffisants,on prendra aussi T3, mais alors il faudra trois axes... D'où un graphique plus compliqué. Je n'ai pas encore bien compris comment on peut en déduire un graphique sur papier, c'est à dire à deux dimensions...

Il me semble utile ici, de montrer un exemple concret pour que des lecteurs naifs en géographie se rendent compte de l'évolution de cette discipline. Cet exemple est extrait d'un ouvrage en préparation, rédigé par H. Chamussy, Maître de conférence à l'Institut de Géographie Alpine, et qui soutient activement la mise en place des stages mathématiques-géographie.

COMMENT CARACTERISER DIVERSES FORMES DE SOUS-DEVELOPPEMENT EN AFRIQUE DU NORD- OUEST.

Les données:

16 pays (Algérie,Bénin, Burkina,Cote d'Ivoire,Ghana, Guinée,Libéria, Mali, Maroc, Mauritanie,Niger, Nigéria, Sierra Léone,Togo, Tunisie.)

12 variables:

PNBH : Produit National Brut par habitant.(en 83)

CPNB : Croissance annuelle moyenne du PNBH de 65 à 83.

ESPE : Espérance de vie à la naissance en 83.

PRIM : Part de l'agriculture dans le Produit Intérieur Brut en 83.

DETT : Service de la dette (% du PIB) en 83

CPOP : Croissance démographique (% annuel) entre 73 et 83.

FECO : Indice conjoncturel de Fécondité en 83.

INDU : Part de la population active travaillant dans l'industrie en 81.

URBA : Taux d'urbanisation en 83.

SCOL : Taux de scolarisation des 6-11 ans en 83.

ARME : Part (% du PIB) des dépenses de l'Etat en Défense en 83.

ETAT : Part (% du PIB) des dépenses du secteur public en 83.

(Sources :Banque mondiale, L'ETAT DU MONDE 1985. La Découverte)

Comme on peut le deviner, avec un tableau de 16 par 12, nous sommes loin de l'exemple 3 par 4 au point de vue des calculs. Mais le principe reste le même.

Tableau 5.1 - Les données

	PNBH	CPNB	ESPE	PRIM	DETT	CPOP	FECO	INDU	URBA	SCOL	ARME	ETAT
Maro	760	2,9	58	17	8,3	2,6	5,8	21	43	49	9,0	20
Alge	2300	3,6	60	6	9,3	2,1	7,0	25	46	83	1,9	16
Tuni	1290	5,0	61	14	7,4	2,5	4,9	32	54	82	1,5	17
Burk	180	1,4	44	41	1,3	1,9	6,5	13	11	13	2,5	14
Maur	480	0,3	44	34	5,0	2,2	6,0	8	25	24	8,7	23
Mali	160	1,2	42	46	1,3	2,5	6,5	12	19	20	3,3	27
Nige	240	-1,2	43	33	5,6	3,0	7,0	3	14	20	0,7	10
Sene	440	-0,5	43	21	1,9	2,8	6,6	10	34	37	2,5	19
Guin	300	1,1	40	38	4,0	2,0	6,0	11	26	28	4,9	19
Libe	480	0,8	49	36	3,2	3,3	6,9	14	38	40	6,6	23
Sier	330	1,1	44	32	0,9	2,1	6,5	19	23	35	1,9	7
Mria	770	3,2	50	26	3,1	2,7	6,9	19	22	75	2,0	11
Beni	290	1,0	43	40	2,5	2,8	6,5	16	16	30	1,8	12
Ghan	310	-2,1	52	53	1,9	3,1	7,0	20	38	60	1,0	6
Civo	710	1,0	47	27	12,9	4,6	6,6	4	44	60	0,9	17
Togo	230	1,1	49	22	6,3	2,6	6,5	13	22	78	1,3	17

Ces données (16 pays, 12 variables) résultent de choix:

- Les variables indiquent toutes des niveaux et des formes de développement;
- les états trop peu peuplés ont été éliminés (Gambie, Guinée Bissau);
- les variables sont toutes exprimées en valeurs relatives, de manière à représenter des structures et non des tailles; elles sont sélectionnées : la variable taux d'inflation a été éliminée car un seul pays, le Ghana, se différenciait nettement des autres; la variable consommation d'énergie a été éliminée parce qu'elle était trop fortement corrélée (r=0,91) au PNB/habitant et seule une des deux variables a été conservée. En ce qui concerne les taux de population active par secteur d'activité, seul celui du secteur industriel a été retenu, car les autres (agriculture, services) en dépendent statistiquement.

Les données sont centrées-réduites pour que chaque variable porte a priori la même quantité d'information.

° La matrice de corrélation (tabl. 5.2)

La matrice de corrélations peut se représenter sous forme d' un graphe (fig. 5.5). Ce schéma présente l'ensemble des plus fortes relations; on distingue:
- un groupe de variables bien corrélées entre elles: le P.N.B. par habitant (PNBH) et l'espérance de vie (ESPE), avec chacun 5 "arêtes", sont au coeur de ce groupe qui s'oppose à PRIM (part de l'agriculture dans le P.I.B);
- ce groupe a aussi une périphérie: la croissance démographique (CPOP), le taux d'actifs dans l'indistrie (INDU) et le taux d'urbanisation (URBA) s'individualisent un peu;
- deux variables forment un groupe distinct, bien moins imposant: dépenses militaires (ARME) et dépenses du secteur public (ETAT);
- La fécondité (FECO) n'est bien corrélée à aucune autre variable.

Ces constatations se retrouvent sur les trois premiers axes factoriels.

° La hiérarchie des facteurs (tabl. 5.3)

Les trois premiers facteurs contiennent près des trois quarts de l'information originelle. Le quatrième semble peu intéressant puisqu'il contient moins d'information qu'une variable (ici, 100/p, soit 8,3 %).

Tableau 5.3 - L'information prise en compte par les facteurs

Facteur	%	% cumulé
1	45,0	45,0
2	19,0	64,0
3	14,7	78,7
4	6,1	84,8

Cet extrait communiqué par H. Chamussy fait partie du livre désormais réedité et renouvelé : " INITIATION AUX PRATIQUES STATISTIQUES EN GEOGRAPHIE" par le groupe Chadule, chez Masson.

V — FAIRE AVOUER LES CHIFFRES

Ce chapitre est issu de la collaboration avec M. Ruffieux. La rédaction complète de notre travail commun a été publiée par le Rectorat de Grenoble (Evaluation des résultats de 85) et envoyée dans les établissements concernés.

Voici un exemple de traitement de l'information numérique par le graphique. L'information est stockée dans les données,mais elle n'est pas apparente.

Moi qui observais des graphiques depuis plusieurs années, je me trouvais pour la première fois en "situation réelle" : la matière première, Maurice Ruffieux, chargé d'Etudes au Rectorat de Grenoble, et soucieux d'une mise en place du suivi des établissements scolaires de l'académie, cette matière première donc, il me proposait de la pétrir avec lui...Une expérience, une chance à ne pas manquer!

(Il est intéressant de noter qu'à l'IREM personne ne comprit qu'il s'agissait là du prolongement naturel et inespéré de ma réflexion sur le numérique...Autrement dit, libre à moi de m'ajouter cette charge de travail, mais ce serait "bénévole", et non estimé comme temps de recherche... Sans doute cette parenthèse est-elle hors sujet, mais la stupidité des conditions de travail d'un enseignant du secondaire est à signaler régulièrement! Tout concourt à nous maintenir dans une routine....qu'on nous reprocherait presque ensuite... Diable, j'ai l'humeur revendicative ce 30 décembre...Peut-être qu'en seconde lecture , ma paire de ciseaux censurera ce paragraphe, mais peut-être aussi que j'oublierai de l'enlever!)

*** 1. Une approche de l'hyperbole.

(...ou " Comment je fis connaissance avec M.Ruffieux, suite aux angoisses de Mr le Proviseur...")

Dans le courant de l'année 84/85, un petit livre rouge arriva du Rectorat . Il s'intéressait aux résultats au bac des divers lycées de l'académie... Le graphique de la page suivante montre l'état des lieux!

Le lycée dans lequel je travaille est situé tout en haut:je l'ai entouré d'un petit nuage qui, bien sûr, ne figurait pas dans le document officiel! Cette situation originale inquiéta notre proviseur. Tout en haut, pourquoi pas, mais en dessous de la sacro sainte moyenne académique... c'était le déshonneur. Nous n'habitons pas à "Ambition sur Isère", mais tout de même!

EVALUATION DES RESULTATS OBTENUS EN 1985

PAR LES LYCEES D'ENSEIGNEMENT GENERAL ET TECHNOLOGIQUE

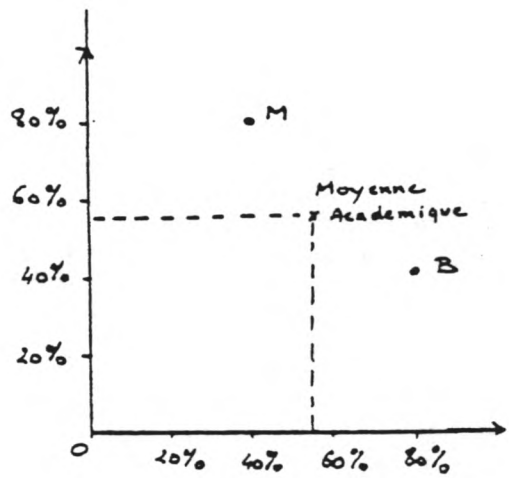
Axe Vertical :

% d'élèves qui, entrés en seconde en 1981-1982 étaient en terminale en 1983-1984

Axe Horizontal : % de ces élèves de terminale qui ont réussi le BAC/BTN/BT en 1984

Pour apprendre à lire ce graphique, prenons un extrait imaginaire:

Axe vertical:
% d'élèves qui,entrés en seconde en 81-82 étaient en terminale en 83-84.



Axe horizontal:
% de ces élèves de term. qui ont réussi le Bac/ BTN / BT en 84

Pour mieux faire comprendre la difficulté de lecture d'un tel graphique,reprenons le procedé de la caricature : B, un "bon" lycée (80% des élèves en terminale en 3 ans ont le bac) et M qu'il faut bien appeler un "mauvais" lycée...

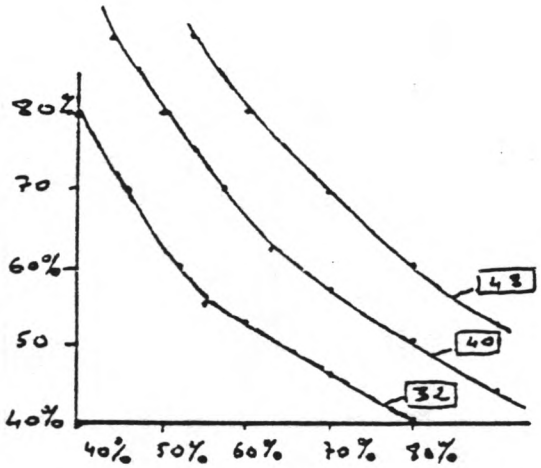
OUI,MAIS...

Sur 100 élèves entrant en seconde au bon lycée, 3 ans plus tard, il n'y en a que 40 en terminale. Le 80 % de réussite porte donc sur ces 40. Il y a donc 32 élèves (40 . 0,80) qui ont le bac 3 ans après leur entrée en seconde...

Et sur 100 élèves entrant en seconde au "mauvais" lycée, 3 ans plus tard il y en a 80 en terminale. La maigre réussite (40%) porte donc sur ces 80. Donc il y en a aussi 32 (80 . 0,40) qui ont le bac en trois ans...

Quant à la moyenne académique,elle donne le même résultat! (56,5 . 0,56 = 31,6)

En fait , les lycées ayant une égale réussite relativement à la population entrée en seconde trois ans plus tôt, se disposent suivant des courbes (ce sont des hyperboles!) qui s'empilent les unes sur les autres. De même qu'on peut dessiner les courbes d'égale altitude (ou courbe de niveau) sur une carte, de même,pour mieux situer les lycées, on aurait pu dessiner les courbes d'égales réussites...

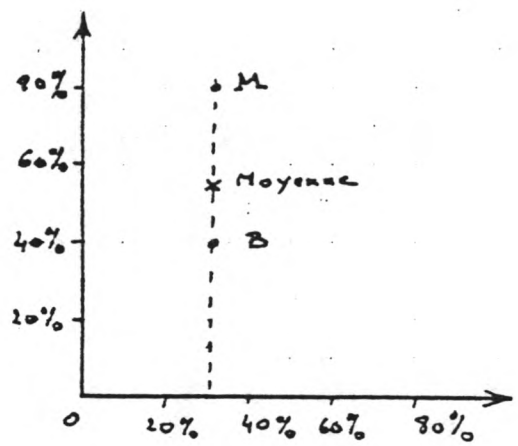


Pour tous les lycées disposés sur cette courbe: Sur 100 élèves entrés en secondes, 48 ont le bac 3 ans plus tard.

Travailler cet exemple en classe permet de faire fonctionner à nouveau les produits de pourcentage. Et les courbes "d'égale réussite sont un bel exemple de produit constant, ce qui n'est pas si fréquent. On passera ensuite à des exemples détachés d'un exemple (Dans un repère placer des points tels que xy = 12) pour faire découvrir la seconde branche de l'hyperbole... l'idée de symétrie est alors immédiate,on peut aussi en profiter...On rencontrera plusieurs fois des cas où un exemple réel n'est qu'une partie d'une courbe mathématique (Deux exemples figurent dans Liaison Mathématiques-Economie). On pourra aussi affiner la notion d'inverse proportionnalité, la faire percevoir comme un cas particulier précis de décroissance. Bref, si ce graphique n'avait pas existé, il aurait fallu l'inventer!

Bien sûr un tel graphique était un peu difficile à lire, aussi désormais, c'est promis, le bureau d'études du rectorat fabriquera des graphiques qui se lisent "en suivant les parallèles aux axes":

Pour cela il suffit de changer la définition de l'axe horizontal:



Axe horizontal:

% d'élèves ayant eu le bac en 3 ans par rapport à l'ensemble des élèves entrés en seconde.

*** 2.Faire "avouer" les données numériques.

(...ou "Que faire quand les nuages -de points- s'amoncèlent dans un bureau du Rectorat"...)

Le nuage de la cohorte 81 fut donc redressé,en voici une simplification pour mieux comprendre la suite.

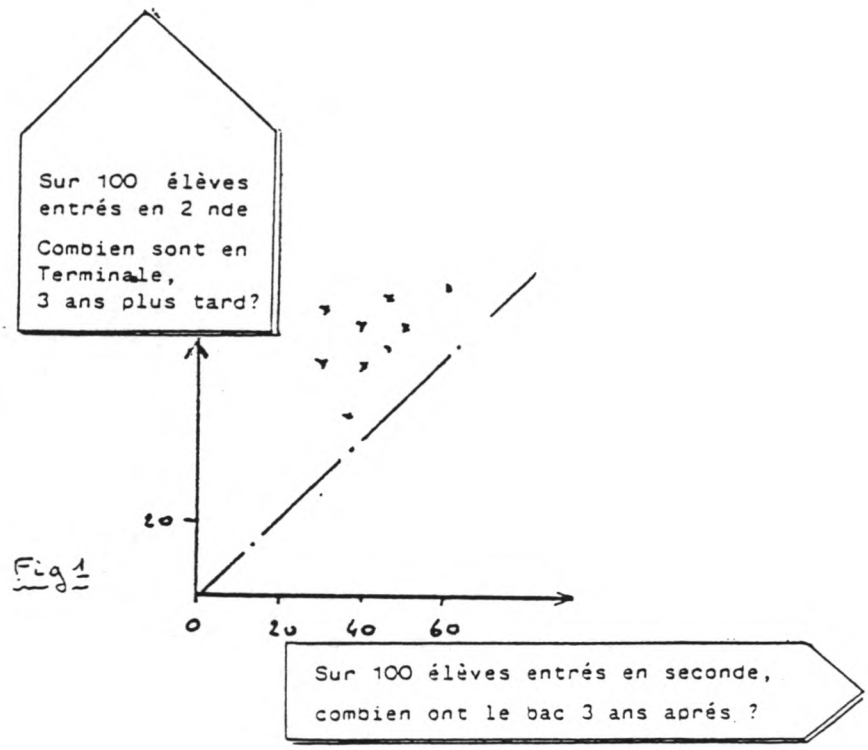


Fig 1

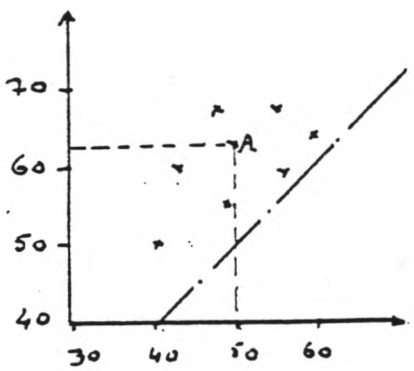
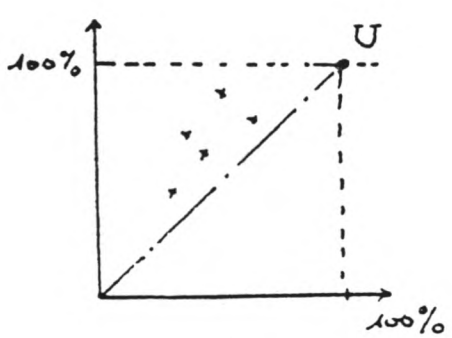


Figure 2

Comme il ne peut y avoir plus d'élèves ayant le bac trois ans après l'entrée en seconde qu'il n'y a d'élèves arrivés en terminale, le pourcentage porté sur la verticale sera toujours plus grand que le pourcentage porté sur l'horizontale. De ce fait,les points représentant un lycée seront tous situés au dessus de la "diagonale" dessinée.(Mais si pour des raisons de cadrage des points sur le papier, les axes ne commencent pas à la même graduation, cette diagonale est décalée. Fig 2)

Exemple: Au lycée A,sur 100 élèves entrés en seconde l'année 81, 63 étaient en terminale trois ans plus tard et 50 ont eu le bac.

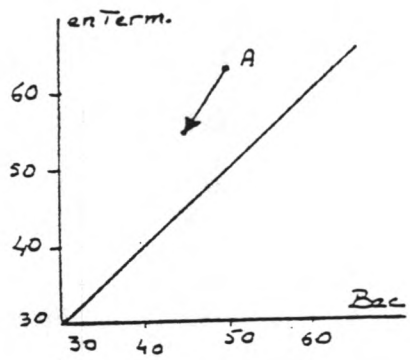


L'idéal (on peut rêver...) serait bien sûr d'avoir 100 % d'élèves en Terminale et tous reçus!

Le lycée Utopie serait donc situé au point " U ".

Puis arriva le nuage de la cohorte 82.

Et avec lui une certaine perplexité...Que faire face à tous ces chiffres traduits en points? Comme de plus nous étions convaincus que la froide comparaison des établissements n'a guère de sens, compte tenu de tous les paramètres qui entrent en jeu, il fallait privilégier le suivi dans le temps de chaque établissement. Nous nous sommes donc lancés dans la fabrication d'un troisième graphique qui serait celui de l'Evolution d'une année sur l'autre.



Exemole:

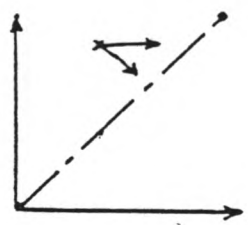
Pour la première promotion des secondes indifférenciées, le lycée A avait:

63% en Term et 50% bacheliers.

Avec la promotion suivante, il a:

55 % en Term. et 45% bacheliers.

Mais que signifierons les flèches obtenues? Seront-elles porteuses de sens? (pour une flèche, ce serait la moindre des choses!) Nous avons alors simulé des flèches possibles:



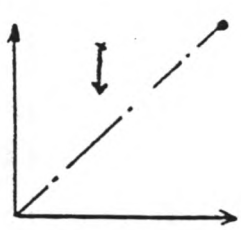
* s'approcher du point "idéal"?

- a l'horizontale?

Mais quel miracle pourrait amener ce progrès absolu? Les élèves étant globalement les mêmes, la "sélectivité" restant constante (déplacement horizontal), s'agit-il d'un profond changement de pratiques pédagogiques?

- en descendant vers la droite?

Si on mène moins de monde en 3 ans en terminale, comment pourrait-il y avoir sensiblement plus d'élèves ayant le bac en 3 ans ?



* s'approcher de la bissectrice?

C'est une tendance naturelle afin de diminuer les échecs au bac (et de faire monter le taux de succès!) Mais pour opérer cette descente, il faut être sûr de faire redoubler à bon escient, c'est à dire de ne pas faire doubler des élèves susceptibles d'avoir le bac en 3 ans...

* le risque est la descente vers la gauche:

Plus de sélectivité,et finalement moins de reçus en 3 ans. Un certain nombre d'élèves ont sans doute redoublé alors qu'ils auraient pu avoir le bac en 3 ans...

Et puis, on se mit à fabriquer les flèches réelles...non sans quelques séances de découragement (ça va être inextricable! un vrai jeu de mikado! illisible! on perd son temps... Qui a eu une idée pareille! ...)

Mais plus on mettait de flèches, plus l'information émergeait de ses brumes numériques. Il est vrai que c'est une curieuse expérience de travailler "pour de vrai" sur des chiffres, sans savoir ce qu'on va trouver...et même si on va trouver quelque chose.

"Est-ce que tu vois la même chose que moi?" fut à peu de chose près notre réaction commune. C'est qu'en effet presque toutes les flèches partaient en bas vers la gauche, c'est-à-dire comme la troisième hypothèse envisagée presque gratuitement...

La tendance générale des 70 lycées de l'académie affleurait là sur le graphique. Ce n'était pas une "moyenne" qui brasse parfois des données trés dispersées, c'était la visualisation du comportement individuel de chaque établissement qui montrait une attitude curieusement convergente. Lycée des villes, lycée des champs, tous le même recul aprés la cohorte 81 qui avait usé du droit de passage automatique de première à Terminale, lequel droit avait été aboli avant qu'on puisse en mesurer les conséquences...

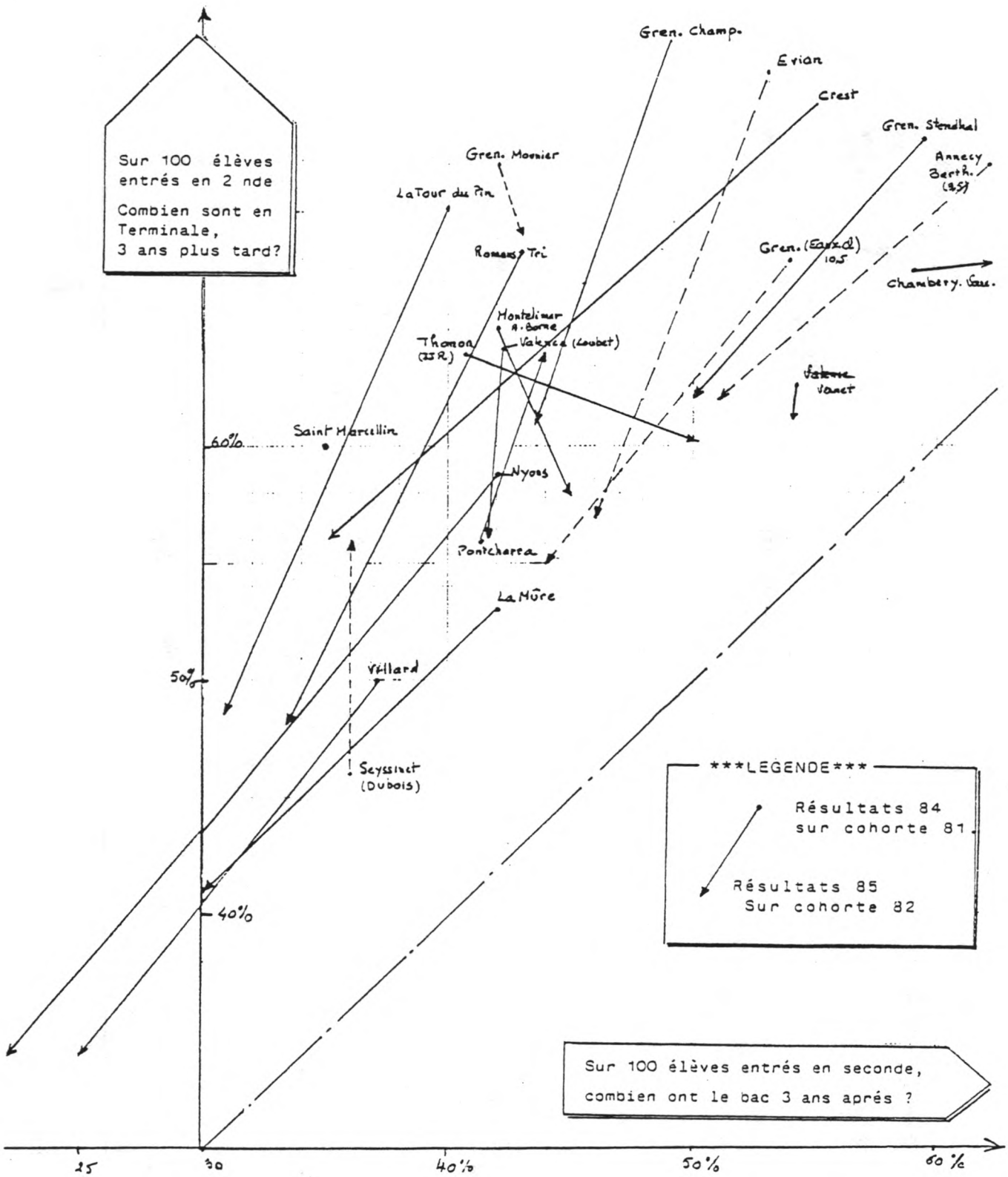
A la page suivante figure un extrait des 70 flèches. C'est le groupe des lycées "classiques".

Notre sujet n'est pas ici de commenter ce graphique,mais de montrer comment il peut faire cristalliser une information dissoute dans les données numériques...D'autres données attendent encore qu'on les fasse parler (en particulier la réussite en 4 ans de lycée par rapport à la réussite en 3 ans ...Nos essais de graphique nous ont laissés dans l'expectative. Alors on continue à imaginer comment on pourrait torturer graphiquement ces chiffres pour qu'ils nous disent ce qu'ils savent...)

Bien sûr aussi, ces graphiques sont difficiles, ce ne sont pas des images de l'information pour la rendre plus digeste. Il faut apprendre à les lire. Il faut les livrer avec un "mode d'emploi" afin de ne pas imposer une "preuve" inaccessible.

I *** GROUPE DES LYCEES ENVOYANT MOINS DE 6% D'ELEVES EN TT et TI

(sections tertiaires et industrielles)



VI - DEMYSTIFIER LE CHIFFRE

Le grand galop des " pour cent" déferle désormais dans nos informations quotidiennes, cravache nos cervelles trop lentes, talonne notre inconcevable lourdeur d'esprit et nous nous retrouvons, bien écoutant mais mal comprenant, face à nos célèbres jockeys politiques faisant cavaliers seuls dans des raisonnements numériques qui nous laissent démunis, gênés, "malaisés", bref, honteux....

" Si j'y tenais mal mon rôle, c'était de n'y comprendre rien" dit un poème d'Aragon. Ainsi en est-il souvent de l'enseignant face à ce nouveau langage.

En guise de conclusion, je rapporterai ici une partie d'un travail collectif mené avec P.Frappat (Sciences Economiques) et R.Samson (Philosophie) et les élèves d'une Terminale B.

Notre idée était d'apporter nos divers éclairages de spécialiste d'une discipline pour tâcher d'apporter quelques éléments de réponses à la question: la connaissance quantitative est-elle automatiquement scientifique? objective ?

- pourquoi y a-t-il eu, et y a-t-il encore aujourd'hui un fétichisme du chiffre?

- Quels sont les différents pièges de la connaissance quantitative? Pourquoi est-elle devenue un obstacle épistémologique (voir Bachelard: " La formation de l'esprit scientifique").

je découvris à cette occasion que Pythagore n'était pas qu'un théorème:

§ " Toute chose qui existe possède un nombre,car il n'est
§ pas possible que quoi que ce soit puisse être connu ou
§ même imaginé sans nombre" dit un disciple de Pythagore.
§ (cité dans "La perversion mathématique" de A. Upinski.)

Platon et Euclide m'apparurent aussi sous un autre jour, mais je n'ai pas assez de recul sur ces sujets historiques et philosophiques pour pouvoir en écrire...

je me limiterai donc aux thèmes ayant un rapport plus direct avec le présent fascicule, à savoir l'actualité de l'information numérique. (Nous avons assez peu abordé ce qui touche aux sciences expérimentales : méconnaissance de l'importance de l'instrument de mesure, excès de précision, méconnaissance de l'ordre de grandeur, puisque les sections B n'ont ni physique, ni biologie en terminale).

*** Y a-t-il un double langage des chiffres?

A cette question, un livre intitulé "La perversion mathématique" répond OUI avec insistance,et Renée, notre philosophe, en est quelque peu ébranlée, d'autant plus que l'auteur est présenté sur la couverture, comme étant le "chef de file de la nouvelle génération de mathématiciens". Je signale, au cas où votre ignorance impardonnable ne vous aurait pas permis de l'identifier, qu'il s'agit de A. Upinsky. A la page 31 (aux éditions Rocher), l'auteur cite d'anciennes statistiques exhumées par Jean Fourastié ("Idées majeures",éditions Gonthier) qui peuvent prêter à sourire:

En effet, dans les années 1730, l'espérance de vie à la naissance était d'à peine 25 ans, mais l'âge moyen au moment du mariage était de 25 ans pour les femmes... et 27 pour les hommes, ce qui n'empêchait pas la durée moyenne du mariage d'être de 17 ans!

Devant cet apparent paradoxe...que des élèves de seconde ont su démonter rapidement (bien que leur "prof de math" ne soit pas chef de file d'une génération de mathématiciens!), Upinsky,lui , croit prouver ainsi que "tel est le double langage des nombres". (page 31).

En fait cet exemple est à rapprocher de l'exemple sur "le revenu moyen des agriculteurs qui augmente"... En effet, la déduction la plus naturelle, celle d'un lecteur non formé au langage des chiffres, aboutit à ce qui semble être une plaisanterie: peut-on se marier en moyenne deux ans après son décès?

C'est bien sûr une forte mortalité infantile qui fait baisser l'espérance de vie. Mais pour qui arrivait à franchir ce cap difficile, il était possible de convoler en juste noce de son "bien-vivant"!

Ce qui me paraît grave ici, c'est la nullité de l'argument joint à la prétention du titre associé (mathématicien!) pour faire passer le tout.

Pour répondre à de tels arguments, je prendrai une comparaison : si un chinois me propose une affirmation dans sa langue maternelle je proclamerai ma simple ignorance de cette langue, et je n'en déduirai pas, comme un fait dûment établi, le double langage du chinois!

J'ai choisi de m'arrêter sur cette notion de moyenne, car c'est une notion très répandue. Elle nous est familière depuis qu'en notre âge d'enfance scolaire nous la surveillions jalousement, donc on ne s'en méfie pas.

Nous avons déjà montré la subtilité du langage numérique, nous avons même fait sentir que parfois, la moyenne de l'école primaire (on ajoute tout et on divise...) n'est pas légitime et qu'il faut pondèrer. Mais même quand la moyenne arithmétique est appropriée à la situation, elle doit être assortie de commentaires, comme devraient l'être d'ailleurs tous les chiffres cités.

Voici par exemple comment M.Ruffieux et moi-même, avons accompagné l'énoncé de la "moyenne académique" (de la réussite au bac d'une cohorte d'élèves), dans la brochure sur l'évaluation des résultats au bac:

EVALUATION DES RESULTATS OBTENUS EN 1985

*** PARLONS " MOYENNE ".

La moyenne est le plus usuel des paramètres statistiques: quand l'information est trop vaste,on essaie de la résumer . On peut toujours calculer une moyenne,mais celle-ci a plus ou moins de sens.

Ainsi, il faisait , hier 28 mai, une température de 26° et aujourd'hui 14°. Que signifie la moyenne de 20°? Me réchauffe-t-elle ce jour?

On peut toujours faire le calcul d'une moyenne.Mais l'exactitude mathématique de ce calcul n'infère rien quant à la représentativité de celle-ci.

En gros, plus les résultats sont dispersés,moins la moyenne a de sens...ce que tout le monde pressent sans avoir besoin des sciences mathématiques! Une moyenne devrait donc toujours être accompagnée d'un indicateur de la dispersion des résultats,ce qui est ici le cas des graphiques 2 et 3.

Ainsi,sur cet exemple simple, pouvons nous apercevoir l'une des causes du malentendu statistique entre ceux qui les élaborent et ceux qui subissent les conclusions. Lorsqu'une moyenne est "refusée" instinctivement par des lecteurs, c'est peut-être qu'ils refusent de voir une réalité, mais c'est peut-être aussi que la moyenne est dépourvue de sens pour cause de trop grande dispersion...Le calcul peut être légal et le rejet du résultat...légitime!

(Disons aussi que les vrais statisticiens sont en général trés prudents et que leurs résultats sont souvent une batterie de chiffres et non un chiffre unique, le tout assorti de commentaires.Ce sont souvent les intermédiaires, ceux qui veulent communiquer,transmettre les résultats qui isolent un résultat de son contexte.Ce peut être pour faire spectaculaire,mais aussi pour faire simple...et être mieux reçu du lecteur!)

Le langage numérique est un langage difficile, on risque souvent le faux-sens ou le contre-sens.Mais se heurter à des difficultés, n'est-ce pas aussi former sa réflexion et son sens critique?

Ce qui est grave quand les commentaires sont insuffisants, c'est qu'un lecteur non initié ne sent même pas qu'il tombe dans des pièges. Il en va donc de l'honnêteté de la communication : il faut prévenir par un "Attention, danger!"

Puisque le langage numérique est difficile, et donc mal compris, comment expliquer son succès? Nous distinguerons des causes secondaires, et une cause principale: l'impact du chiffre.

*** Quelques causes possibles pour expliquer la prolifération des chiffres dans nos informations.

.a) Dans les pays industrialisés, la collecte des données numériques ainsi que leur traitement sont devenus aisés. On pense en premier lieu à l'ordinateur, mais on oublie que la collecte des chiffres repose aussi sur une administration développée et centralisatrice. Ainsi les pays en voie de développement sont-ils souvent démunis...même avec l'informatique. Comment en effet étudier la corrélation entre l'age de la mère au moment de l'accouchement et la mortalité infantile si les femmes ne savent pas leur age exact faute de registre d'état civil à l'époque de leur propre naissance?

Il est intéressant de noter que cette pauvreté des données dans certains pays a fait naître, dans les années 55, une nouvelle discipline : " L'analyse des données imparfaites en démographie". (Voir l'article de Rémy Clairin, dans STATECO,revue de l'INSEE, n° 43)

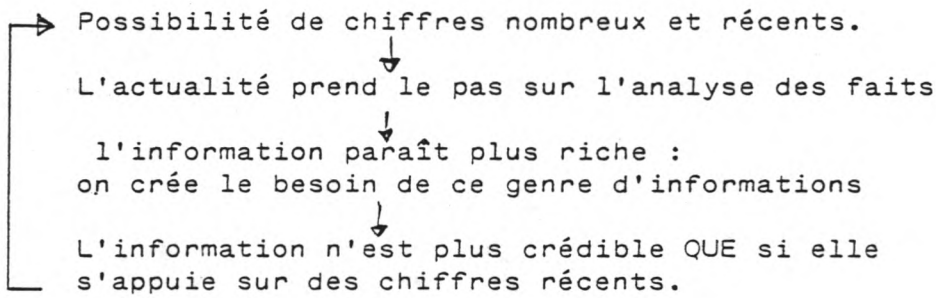
Pour ce qui est des pays industrialisés, le problème n'est pas de trouver des chiffres, mais de trier devant l'abondance de la récolte! Aprés l'âge de pierre et l'âge de bronze, ne sommes-nous pas actuellement en train de vivre le sond'âge?

.b) En une génération, le vocabulaire économique de l'auditeur ou du lecteur moyen a nettement progressé. Inflation, Pouvoir d'achat, par exemple, font désormais partie du langage courant,ce qui ne signifie pas d'ailleurs qu'on saurait aisément les définir, mais ils sont compris globalement. L'information ou l'explication qui pouvait s'abriter derrière ces termes quand ils faisaient "savants", ne le peut plus.

Il en est sans doute de même pour la géographie, la biologie dans le monde scolaire. Ce que le mathématicien Yves Chevallard propose dans son livre "La transposition didactique" trouve ici un soutien. Celui-ci émet l'hypothèse que les savoirs enseignés par l'institution doivent paraître "savants" aux yeux des parents afin que l'école, et les enseignants, aient une autorité légitimement reconnue par les consommateurs d'école que sont les parents (par enfants interposés bien sûr). Le langage des chiffres encore aussi obscur que les hiéroglyphes avant Champollion, peut remplir cet office. Dans les manuels de géographie, cela est spectaculaire, le "sida numérique" a tout contaminé...

.c) Une baisse de crédibilité de la parole du "spécialiste" si celle-ci s'appuie seulement sur un travail antérieur "mystérieux". L'apport de chiffres donne l'illusion de la transparence et renforce l'actualité du discours, donc sa force de preuve au moment où il parle. Un chiffre plus récent devient un chiffre plus vrai, la course à la superposition des sondages nous le prouve. La qualité du chiffre n'est jamais mise en cause (comment est-il obtenu, méthodes? critères?) mmais seulement son degré d'actualité.

La primauté de l'Actuel est-elle une cause de cette invasion numérique ou bien une conséquence de l'accession de plus en plus aisée aux chiffres? Peut-être est-ce une cause récurrente qui permet au système de s'emballer:



.d) Une nécéssité économique.

Prendre une décision, pour des responsables d'entreprise, ne peut plus guère se faire...à l'intuition. " les entreprises attendent impatiemment les statistiques économiques préliminaires essentielles, telles que les chiffres hebdomadaires concernant la masse monétaire et les estimations trimestrielles du PNB" (Article de John Keane, dans la revue STATECO,déja citée)

*** Le nombre est d'or!

En cette fin du second millénaire, la magie du chiffre opère toujours et le fait que les chiffres ne soient plus une denrée rare ou précieuse n'a en rien diminué leur pouvoir.

Dans un discours, le chiffre en lui-même est devenu une preuve. Qu'il n'apporte pas grand chose, ou même qu'il ne veuille rien dire, le chiffre est là, inexpugnable soutien du raisonnement obligé. En effet l'usage actuel du chiffre est bien plus d'IMPOSER que de convaincre.

Pourtant le chiffre s'impose car il arrive auréolé de sérieux, de rigueur, d'objectivité, en un mot de vérité : il est l'information "scientifique".

En fait , s'il y a de la rigueur, elle s'exerce dans les opérations que l'on fait avec les chiffres, mais si les données sont douteuses, quelque soit l'exactitude des opérations, les résultats le seront tout autant! Or tout se passe comme si le traitement numérique avait le pouvoir de blanchir les données...

L'an passé, une série d'articles parus dans le Figaro magazine en était la vivante illustration. Partant de taux de fécondité hypothétique (puisqu'on les suppose constants sur 20 ans), le journal annonçait des résultats indiscutables. Et comme il y eut certaines réactions,il y eut des réponses:

" la population est parfaitement connue sous réserve d'une triple incertitude"! écrit G.F. Dumont dans le numéro du 2 novembre 85. (Imaginer un candidat au bac qui annoncerait une connaissance parfaite associée à une triple incertitude... Que veulent dire les mots?)

" Nous avons utilisés les trois données suivantes:
- hypothèse de mortalité,
- hypothèse de fécondité...
etc"

Les données sont des hypothèses...

Et la confusion voulue :
Le journaliste "a commenté l'une des hypothèses mathématiques qui montre ce qui pourrait se passer si la France s'enfermait dans le malthusianisme" (Dumont; 9 novembre 85)

*** Peut-on tout mesurer ?

Evoquons rapidement le Quotient Intellectuel dont l'inventeur, Binet , ne prévoyait pas le succès qu'il allait remporter...pour devenir finalement une sorte de mesure de l'intelligence. Comme le commun des mortels ne sait pas tres bien ce que mesurent les tests de Binet, ni tres bien non plus, et pour cause, ce qu'est l'intelligence, cela fonctionne... On fait du précis dans le plus grand flou.

Mais qui mesurera les conséquences du changement intérieur, souvent inconscient, des parents à l'égard de l'enfant dont ils viennent d'apprendre, de la bouche d'un Spécialiste (avec un grand S!) que son quotient intellectuel est 130 ou 80 alors que la mesure moyenne est 100 ? Heureusement, le plus souvent les psychologues se contentent de situer l'enfant par rapport à un âge : il a 5 ans, mais, actuellement son age mental est de 6 ans 4 mois... ou 4 ans 2 mois. Cela parait déja moins absolu, une avance ou un retard par rapport à un peloton moyen, chacun sait que cela n'est pas nécéssairement définitif.

Voici un autre exemple d'usage négatif du chiffre, cité par Albert Jacquard dans " Au péril de la Science?" (Seuil. Points Sciences. page 37).

Dans les années 30,le docteur Réné Martial, chargé de cours à la faculté de médecine de Paris, propose dans son ouvrage "La race française" un exemple d'indice inventé par lui,"l'indice biochimique du sang". Cet indice est basé sur les fréquences d'apparition des divers groupes sanguins A,B,AB et O, lesquelles fréquences sont assez variables en Europe.

La formule de l'indice biochimique du sang fait on ne peut plus scientifique:

$$I = \frac{f_A + f_{AB}}{f_B + f_{AB}}$$

Pourquoi pas?... Or il se trouve que cet indice est voisin de 3 pour les Français et les Allemands, mais 1,6 pour les Juifs, 1,2 pour les Polonais, et 0,9 pour les "Nègres"...

> *" ....Il a suffi d'imaginer un indice,calculé de façon assez compliquée pour dérouter les esprits non prévenus, et d'utiliser cet indice comme une échelle de valeur.*
>
> *Car il est exact que "l'indice biochimique" des Français est supérieur à celui des Polonais; ce qui est arbitraire est d'en déduire que les Français sont supérieurs aux Polonais; "*

nous dit Albert Jacquard dans son livre dont je ne saurais trop recommander la lecture.

*** Le chiffre pour abattre les préjugés....

Mais A. Jacquard nous montre aussi que des chiffres et du calcul mathématique (du niveau de terminale) peuvent mettre à mal l'idée de famille porteuse de tare dans le cas de maladies héréditaires.

Ainsi certaines familles ont un enfant atteint de mucoviscidose, (un enfant malade sur 2500 naissances) et on pourrait craindre des risques pour la descendance des frères et soeurs... Il est vrai que ce risque existe, mais il existe à l'état latent chez 4% de la population "saine".

Il en est ainsi pour bon nombre de maladies héréditaires. Ainsi, pour la phénylcétonurie (qui entraîne une détérioration des facultés intellectuelles si elle n'est pas soignée à temps), il y a deux cents fois plus de porteurs "sains" que de malades. Une famille sans maladie héréditaire déclarée , ne sait pas en fait ce que contient son patrimoine génétique, d'où le non-sens de certaines mises à l'écart.

Le marteau est-il responsable des conséquences des coups qu'il donne? Il en est de même pour les chiffres qu'il faut considèrer comme des outils, certes plus sophistiqués qu'un marteau, mais des outils quand même.

La connaissance numérique ne s'oppose pas aux autres,elle est une approche parmi d'autres qu'il serait dommage de méconnaître ou de refuser.A l'inverse, elle ne détient pas le monopole de la "vérité", ni la marche supérieure d'un podium sur lequel prendraient place,par ordre "d'objectivité", les divers aspects de la connaissance .

Connaissance quantitative et connaissance qualitative sont en étroite relation, et peut- être a-t-on eu tort de croire que l'on pouvait travailler séparément chaque élément de ce "système". C'est la connaissance qualitative qui donnera de précieuses indications sur la nature des chiffres à collecter. C'est le traitement de cette moisson numérique qui révèlera peut-être une nouvelle facette de la réalité. Cette nouvelle perception , si elle n'est pas refusée à priori, devra être à son tour absorbée par l'approche qualitative en de nouvelles questions, de nouvelles hypothèses, peut-être de nouvelles conclusions ...toujours provisoires. C'est la connaissance qualitative qui permet la naissance de nouvelles interrogations à partir de résultats quantitatifs, et qui de ce fait en demanderont d'autres. Ce sont ces allers-retours qui sont source d'évolution féconde.

Mais grande aussi est la responsabilité de tous ceux qui ont à communiquer la connaissance : enseignants, manuels scolaires, journaux, télévision... c'est là aussi que naît souvent ce fractionnement artificiel des divers miroirs de la connaissance. Peut-être s'agit-il de "faire plus simple" , de "dire pour être compris", mais si on déforme, n'est-ce pas plus grave?

* * * * *

Pour les lecteurs qui ont tenu jusqu'au bout de ces pages parfois arides, je terminerai par un sourire pour nous détendre: nous allons voir comment certains fanatiques de la "mesure à tout prix" opèrent parfois!

L'exemple est extrait du livre "L'intelligence gaspillée" de Michel Schiff,page 56 (SEUIL). L'auteur relate le travail d'un groupe de chercheurs voulant réaliser une comparaison d'aptitudes entre groupes provenant de cultures différentes. L'un des tests était la performance possible sur une bicyclette ergonométrique, engin que les Pygmées africains ont dû découvrir avec étonnement:

" L'idée de faire monter des chasseurs-cueilleurs, qui ne se déplacent qu'à pied, sur une bicyclette est déjà assez distrayante, mais on peut supposer que, malgré quelques problèmes d'apprentissage, ils en comprennent vite l'intérêt. Par contre, imaginer qu'ils puissent, au soleil d'une clairière, déployer tout l'effort physique possible pour faire tourner les pédales d'une bicyclette sans roues qui ne se déplace pas relève de l'urgence psychiatrique. C'est pourtant ce que des spécialistes ont fait non seulement chez les Pygmées mais partout, de la forêt équatoriale aux glaces du pôle, en passant par tous les déserts et les cités de la planète!"

Sans doute vaut-il mieux en rire...que de mesurer notre aptitude à rire!

* * * * *

Je souhaite que ce fascicule vous donne envie de prolonger en puisant dans la bibliographie citée,et je renouvelle mon appel afin de l'enrichir dans l'avenir...

Sylviane GASQUET

Lycée Champollion - IREM
(Grenoble)

FASCICULE LANGAGE DES GRAPHIQUES

S O M M A I R E

*** I . Du croquis à l'organigramme.... . . . . .page 6

Des manuels scolaires à la bande dessinée : bref aperçu des illustrations proposées aux élèves.

*** II . Graphiques et proportionnalité . . . . . .page 11

Quand la proportionnalité se déguise en graphiques: des camemberts à géométrie variable, de la taille des vaches en Europe de l'ESt, du bon usage des toupies, ...et des pyramides du haut desquelles des millions d'immigrés nous contemplent.

*** III . Graphiques gradués. . . . . page 33

Des graduations en tous sens, des axes autant qu'on en veut, des graphiques en peau de chagrin...et des risques de confondre ce que l'on fait en géographie et en mathématiques.

*** IV . Graphiques logarithmiques. . . . . page 54

Le miroir déformant des graphiques logarithmiques: graduations bizarres, impressions fausses, un langage secret...tout cela nous vient du pays des logarithmes dont on vous raconte l'histoire en une brève chronique (Deux pages communes avec le fascicule Langage Numérique)

*** V . DES FICHES ELEVES. . . . . page 74

Quelques exemples de fiches niveau troisième/seconde. Le savoir-faire sur les graphiques de géographie peut-il être récupéré pour réconcilier des élèves avec le cours de math?

* * * PETITE BIBLIOGRAPHIE * * *

* groupe CHADULE (H. Chamussy, J. Charre, P. Dumolard, M.-G. Durand, M. Le Berre) : INITIATION AUX PRATIQUES STATISTIQUES EN GEOGRAPHIE. Edition 87, MASSON.

* CIBOIS Philippe : L'ANALYSE FACTORIELLE . Collection Que sais-je ? Presses Universitaires de France.

* JACQUARD Albert : AU PERIL DE LA SCIENCE ? Collection Points-Sciences Editions du SEUIL.

* JACQUARD Albert : ELOGE DE LA DIFFERENCE . Editions du Seuil.

* JACQUARD Albert : MOI ET LES AUTRES . Initiation à la génétique. Colection Points- Virgule . Editions du SEUIL.

* KLATZMANN Joseph : ATTENTION STATISTIQUES ! Comment en déjouer les pièges. Cahiers libres 405. Editions LA DECOUVERTE .

* LEVY M.-L., EWENCZYK S. , JAMMES R. : COMPRENDRE L'INFORMATION ECONOMIQUE ET SOCIALE . Guide méthodologoqie. Editions HATIER.

* MEIRIEU Philippe : L'ECOLE MODE D'EMPLOI . Des "méthodes actives à la pédagogie différentiée. Editions ESF .

* LEGER Alain : ENSEIGNANTS DU SECONDAIRES . Collection L'éducateur. PRESSES UNIVERSITAIRES DE FRANCE.

* REUCHLIN Maurice : PRECIS DE STATISTIQUES. Collection Le Psychologue. Presses Universitaires de France.

* SANDERS, MURPH et ENG : LES STATISTIQUES, UNE APPROCHE NOUVELLE. Editions MAC GRAW - HILL .

* SCHIFF Michel : L 'INTELLIGENCE GASPILLEE . Inégalié sociale, injustice scolaire. Editions du SEUIL.

* UPINSKI Arnaud : LA PERVERSION MATHEMATIQUE. Editions ROCHER.

* * * * * *

....Dernière heure....

- Un peu difficile, mais tout à fait dans le sujet (chapitre 7 surtout): Albert Jacquard . L'Héritage de la liberté. SEUIL.

- Les Résultats des Lycées. N° 3. Bac 86 . (M.Ruffieux, S.Gasquet)

  Rectorat de Grenoble.(disponible dans tous les lycées de l'académie de Grenoble.) Des questions posées par des chiffres....