Les histogrammes : quel enseignement au collège et au lycée ?

Éric Roditi

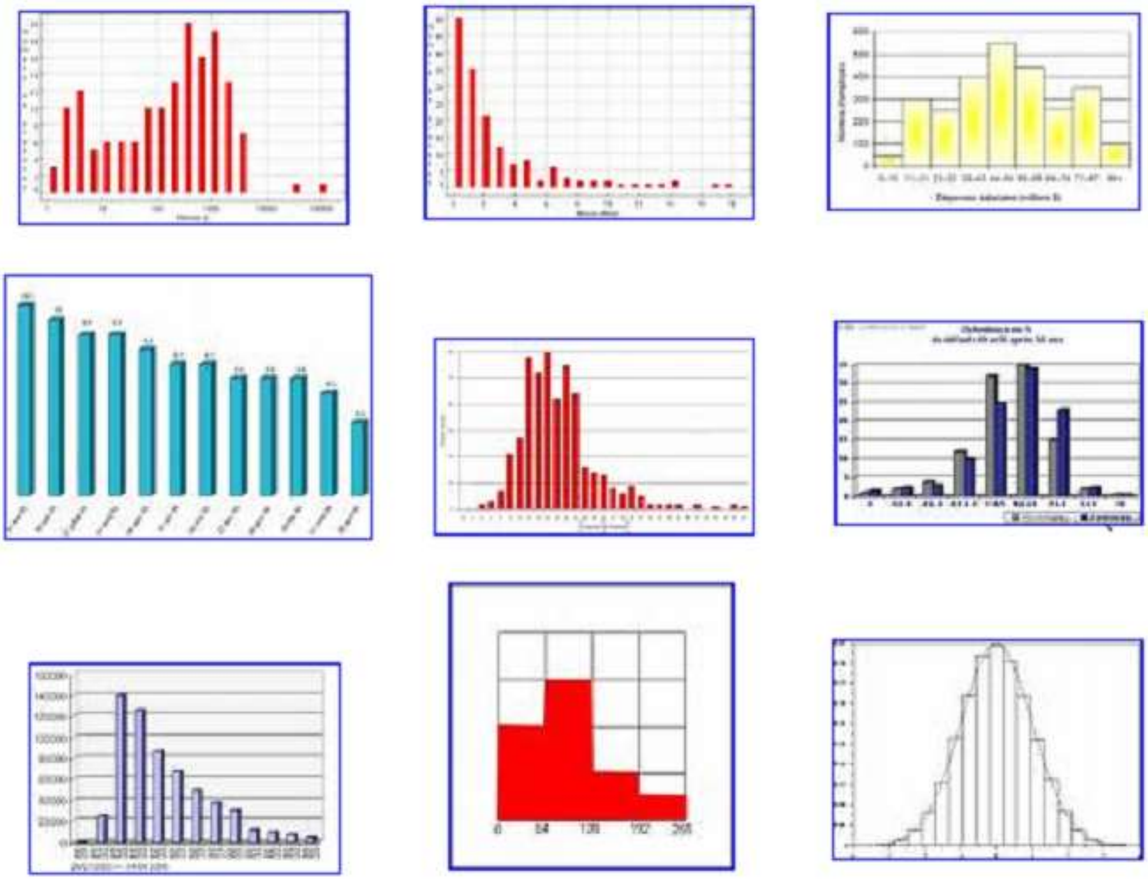
Université Paris Descartes, Laboratoire EDA
Groupe IREM Didactique de l'Université Paris Diderot [20]
CII Didactique

L'enseignement des statistiques vise simultanément la formation mathématique et l'éducation citoyenne. Celui des graphiques, à cause de leur présence importante dans les médias, est très valorisé. Une simple consultation des programmes et des manuels scolaires permet en effet de constater que les diagrammes « en bâtons » ou « à bandes » sont des objets d'enseignement depuis le cycle 3 de l'école élémentaire jusqu'au lycée, c'est-à-dire au premier et au second cycle. Ces diagrammes figurent aussi dans l'enseignement supérieur, et dans les ouvrages, le nombre de pages accordées à la statistique descriptive est d'autant moins important que le niveau mathématique des étudiants est élevé. La place accordée dans l'enseignement à ces diagrammes statistiques que nous qualifierons de « rectangulaires » semble donc suffisante.
Un constat pourtant vient questionner ces premières impressions : les professeurs de mathématiques trouvent généralement peu d'intérêt à cet enseignement, sinon pour saisir une occasion de faire fonctionner la proportionnalité de deux grandeurs : la fréquence (ou l'effectif) d'une valeur et la hauteur du rectangle associé. Une remarque enfin, les programmes du secondaire et les documents d'accompagnement laissent entendre que l'histogramme est un objet mathématique plus difficile à comprendre et à utiliser qu'il n'y paraît, notamment à cause du cas où les classes de valeurs de la variable ne sont pas toutes de la même amplitude. Ce cas n'est d'ailleurs pas au programme du collège, et il n'est à traiter au lycée que dans la circonstance particulière où les classes extrêmes sont les seules à avoir une longueur différente de celle des autres classes.
Les membres du groupe didactique de l'IREM de Paris Diderot (Paris 7) enseignent du collège à l'université, ils ont mené une recherche sur l'histogramme. D'abord sur l'objet mathématique lui-même, sur sa définition comme sur l'utilisation et l'interprétation de ce diagramme. Sur son enseignement ensuite. Cela comprend une interrogation de type écologique sur l'objet de savoir « histogramme » dans les programmes, c'est-à-dire d'une part sur les relations que cet objet de savoir entretient avec les autres savoirs au programme, et d'autre part sur sa fonction d'outil pour résoudre des problèmes liés aux savoirs mathématiques au programme. Cela comprend aussi une étude de son enseignement et de son apprentissage en distinguant sa lecture, son interprétation et sa construction à partir d'une série ou d'une distribution statistique. Cela comprend enfin une analyse des choix possibles pour la programmation de cet enseignement au collège et au lycée.
Le compte rendu de ce travail est organisé en deux parties. Dans la première sont exposés les résultats concernant l'étude de l'objet mathématique histogramme, la seconde traite des questions d'enseignement et d'apprentissage.

1. L'histogramme : quelle représentation graphique ?

Lorsqu'on effectue une requête avec le mot-clef « histogramme » dans une base de donnée d'images en ligne, on peut « apprécier » la diversité des diagrammes qui sont associés à cette dénomination ! Nous avons fait l'expérience avec le moteur de recherche *Google* et nous avons obtenu ces diagrammes dès la première page affichée.

[20] Au sein du groupe Didactique, ont particulièrement contribué à ce travail : Fabienne Cissé, Stéphanie Colin, Christine Mémier et Françoise Pilorge, professeures des académies de Créteil et Paris.



Les diagrammes sont composés de rectangles, contigus ou non, ils comportent un axe des abscisses gradué où les valeurs sont indiquées soit de manière habituelle sous les graduations soit par des intervalles entre les graduations. Un polygone est parfois présenté qui joint les milieux des largeurs supérieures des rectangles. Tous ces diagrammes sont-ils des histogrammes à proprement parler ? Sinon c'est que l'usage courant est différent de l'usage en statistique. Il conviendra alors de s'interroger sur l'intérêt, parmi tous ces diagrammes « rectangulaires », d'en distinguer certains qu'on appelle histogrammes, en mathématiques comme pour la classe de mathématiques.

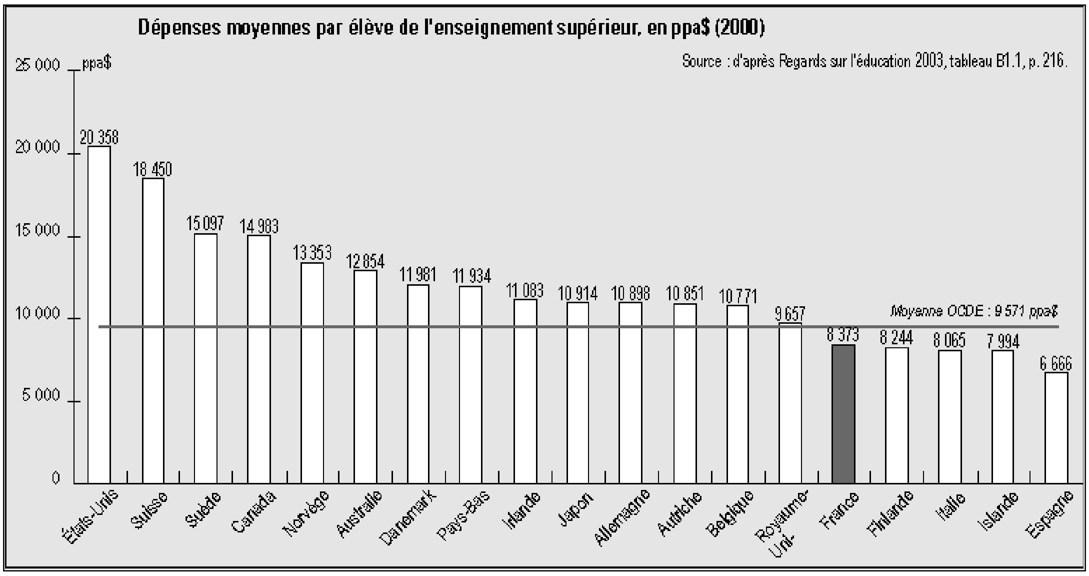
1.1 L'histogramme parmi les différents diagrammes « rectangulaires »

Les critères de classement des diagrammes tiennent aux procédés de représentation des variables représentées. Voici, à partir de la comparaison de cinq exemples de diagrammes « rectangulaires », des critères permettant de reconnaître les histogrammes.

1.1.1. Représentation d'une série ou d'une distribution

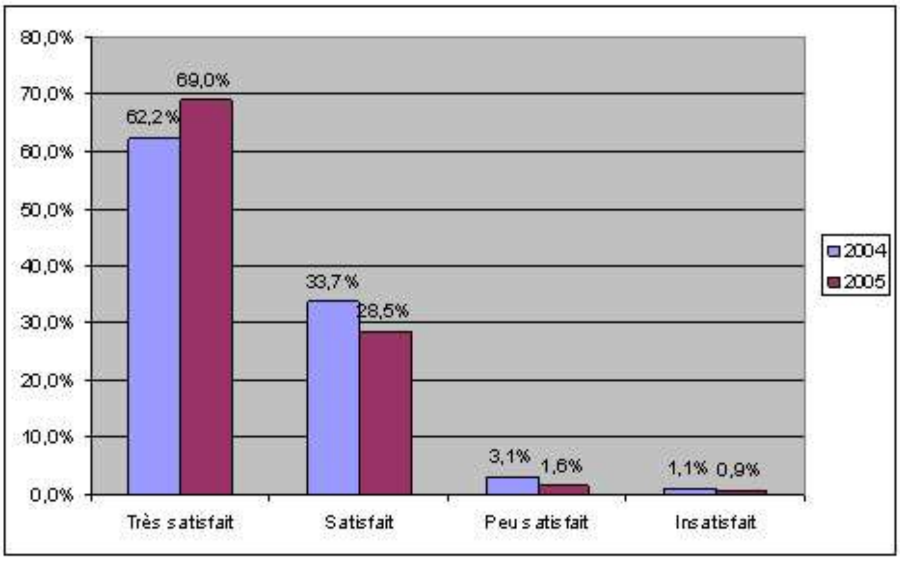
En focalisant sur ce que représentent les diagrammes, on pourrait distinguer ceux dont l'axe des abscisses représente les individus de ceux dont il représente les valeurs de la variable étudiée sur la population.

Ainsi, dans ce diagramme, les individus, en abscisse, sont des pays de l'OCDE auxquels est associée la valeur de la dépense moyenne par élève de l'enseignement supérieur, l'unité étant la parité de pouvoir d'achat en dollars.



De façon courante, puisque les individus ne sont pas ordonnés, ils sont représentés sur le diagramme par ordre croissant ou par ordre décroissant de leur valeur.

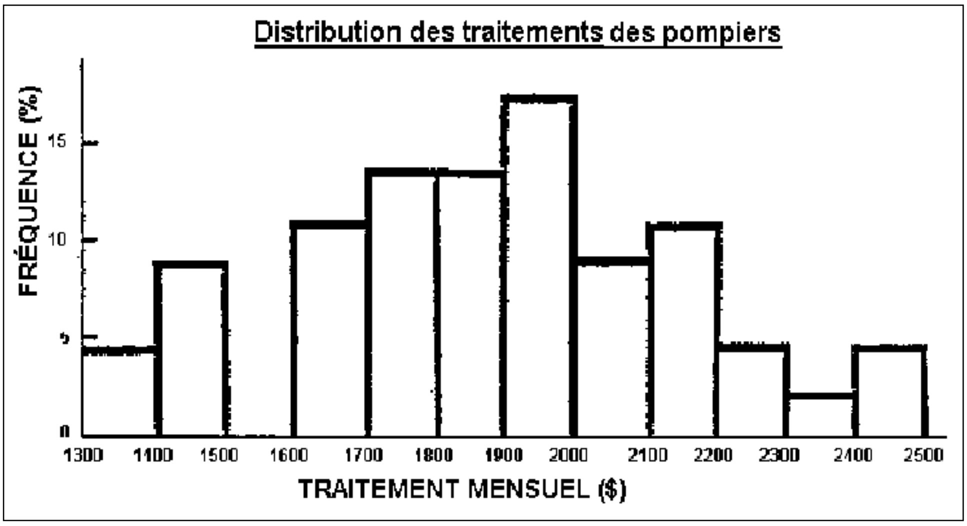
Dans le diagramme suivant, ce sont les valeurs (on dit aussi les modalités de la variable ou du caractère) qui sont en abscisse alors que l'axe des ordonnées indique leur fréquence en pourcentage. Il y a en fait deux diagrammes superposés qui présentent, en 2004 et en 2005, la répartition des valeurs de satisfaction émises par des jeunes après un séjour organisé par l'association « Aventure scientifique ».



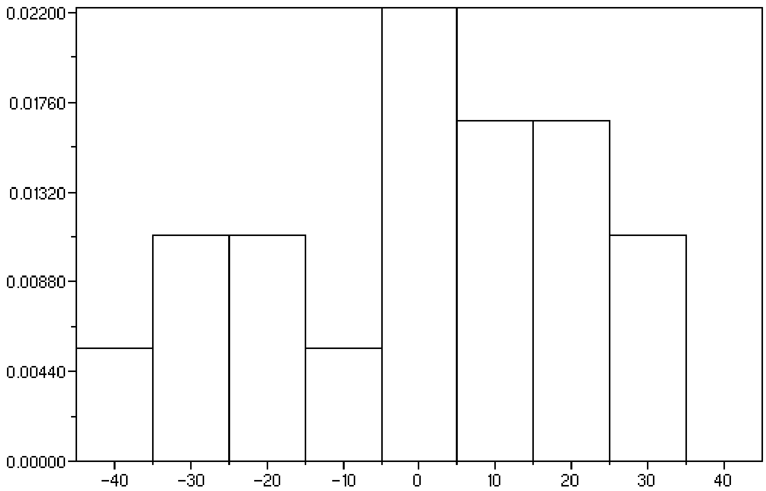
La distinction entre les représentations « rectangulaires » des séries statistiques (tableaux individus – valeurs) ou des distributions statistiques (tableaux valeurs – fréquences) ne nous apparaît pas pertinente pour classer les diagrammes eux-mêmes car, justement, elle ne porte pas sur les diagrammes mais sur les variables représentées. Ainsi, dans un contexte cinématique, une courbe peut représenter la distance parcourue en fonction du temps, une autre courbe peut représenter la vitesse du déplacement en fonction du temps. Et les deux graphiques sont bien des courbes.

1.1.2. Représentation de la fréquence ou de la densité de fréquence

Dans le diagramme précédent, la variable étant ordinale et discrète, les valeurs sont ordonnées sur l'axe des abscisses et les rectangles ne sont pas contigus. Ce n'est pas le cas dans le diagramme suivant qui représente la distribution des salaires mensuels, en dollars canadiens, des pompiers québécois : la variable est dite numérique et continue. Et cela même si certaines valeurs au sein de la classe ne sont pas prises, même si le nombre de valeurs étant nécessairement fini, leur ensemble est par conséquent discret. La variable est qualifiée de continue car, a priori, toutes les valeurs sont considérées comme possibles.



Le dernier exemple représente la distribution des résidus dans une régression linéaire. Pour chaque point du nuage des données, la différence entre son ordonnée et celle de son projeté sur la droite de régression linéaire est un résidu. Dans cet exemple, les résidus sont groupés par classes de longueur 10 à partir de l'origine –45.

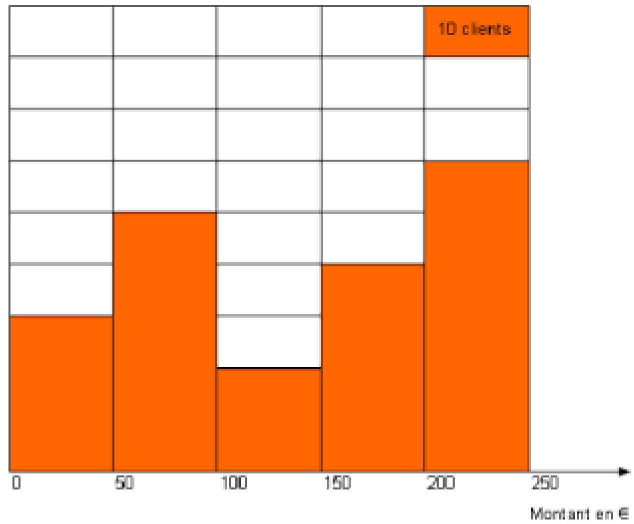


Dans les diagrammes de satisfaction des jeunes comme dans le diagramme des salaires des pompiers québécois, la somme des ordonnées est la somme des fréquences des modalités de la variable, elle est égale à 1 ou 100%. Dans le dernier graphique en revanche, la somme des ordonnées prises par chacune des classes n'est pas égale à 1 mais manifestement à 0,1. Comme la longueur des classes est précisément égale à 10, on en déduit que l'axe des ordonnées n'est pas l'axe des fréquences, mais l'axe des densités de fréquence. D'une manière analogue à la définition de la densité de probabilité, une densité de fréquence se définit comme une fréquence par unité des valeurs de la variable. Sur l'exemple ci-dessus, pour toute valeur comprise entre –5 et 5, la densité de fréquence est 0,022 c'est-à-dire 2,2% par unité de longueur des résidus. Ainsi la fréquence de la classe [–5 ; 5[est égale à $2{,}2 \times 10 = 22\%$.
Ce diagramme diffère fondamentalement des précédents, il s'agit de bien comprendre en quoi. Le diagramme rectangulaire des salaires des pompiers québécois indique que 14% des pompiers ont un salaire compris entre 1 700 et 1 800 dollars. Il n'indique pas que pour chaque salaire compris entre 1 700 et 1 800 dollars, la fréquence est 14%. Et heureusement, sinon en ajoutant les fréquences des salaires 1 700 \$, 1 701 \$, 1 702 \$, etc. on obtiendrait une somme de beaucoup supérieure à 100% ! Ainsi malgré la contiguïté des rectangles, le diagramme des salaires des pompiers québécois est analogue à celui des satisfactions des jeunes après leur séjour de loisirs scientifiques : il s'agit d'un

diagramme en tuyaux d'orgue dont la hauteur indique la fréquence de chaque classe. Ce graphique n'est pas celui d'une fonction dont les valeurs de la variable sont les salaires (variable continue) mais sont les classes (variable discrète). Il en va tout autrement dans le diagramme des résidus, la densité de chaque résidu compris entre –5 et 5 est 2,2 % par unité. Le graphique est celui d'une fonction dont les valeurs de la variable sont tous les résidus. En définitive, le fait que la variable soit continue est en quelque sorte « assumé » par le graphique des résidus, alors qu'il ne l'est pas par le graphique des salaires. Pour terminer ce paragraphe, remarquons que l'aire de la surface totale composée par les rectangles du diagramme est égale à 1 ou, ce qui revient au même, à 100%.

1.1.3. Des diagrammes « rectangulaires » sans axe des ordonnées

Le dernier diagramme, qui permet de faire l'inventaire des types de diagrammes rectangulaires rencontrés, représente la distribution des effectifs des dépenses effectuées sur un site Internet par un échantillon de 200 consommateurs. La variable est continue et les rectangles sont contigus. Ce qui distingue le diagramme présenté de ceux qui ont été étudiés jusqu'à présent est qu'il ne possède pas d'axe des ordonnées. Une légende indique l'aire correspondant à 10 clients.



En forçant un peu notre imagination, nous voyons apparaître un axe des ordonnées à ce graphique : le rectangle de la légende a pour largeur 50 euros (parallèlement à l'axe des abscisses), comme son aire est 10 clients, c'est que sa hauteur (parallèlement à l'axe des ordonnées) est 10 clients pour 50 euros c'est-à-dire 0,2 client par euro. L'axe des ordonnées représente bien une grandeur associée à la variable « montant de la dépense en euros » : sa densité d'effectif. Ainsi finalement, ce cinquième diagramme, malgré l'absence de l'axe des ordonnées, ressemble fort au quatrième : il « assume » la continuité de la variable et il représente la densité. La surface composée par l'ensemble des rectangles a pour aire 200 clients, c'est-à-dire 1 ou 100% si l'on convertit cet effectif total en fréquence totale.

1.1.5 Bilan et définitions d'un histogramme

Pour terminer l'étude de ces diagrammes, nous avons consulté de nombreux ouvrages : dictionnaires, encyclopédies, manuels scolaires et manuels universitaires. Les définitions sont hétérogènes, néanmoins, en cohérence avec nos analyses des exemples de diagrammes, deux catégories de diagrammes rectangulaires se dessinent.

Les diagrammes de la première catégorie sont des diagrammes « à bandes rectangulaires » contiguës ou non, ces bandes rectangulaires s'appellent couramment tuyaux d'orgue, elles pourraient simplement être des bâtons. En effet, le procédé de représentation du diagramme en tuyaux d'orgue repose sur la proportionnalité entre la hauteur du rectangle tuyau et la valeur associée que ce rectangle représente. Ainsi lorsqu'un tel diagramme représente une série statistique (tableau individus – valeurs), chaque rectangle représente un individu et la valeur qui lui est associée, la hauteur du rectangle est proportionnelle à la valeur. Lorsqu'un tel diagramme représente une distribution statistique (tableau valeurs – fréquences), chaque rectangle représente une valeur (ou une classe de

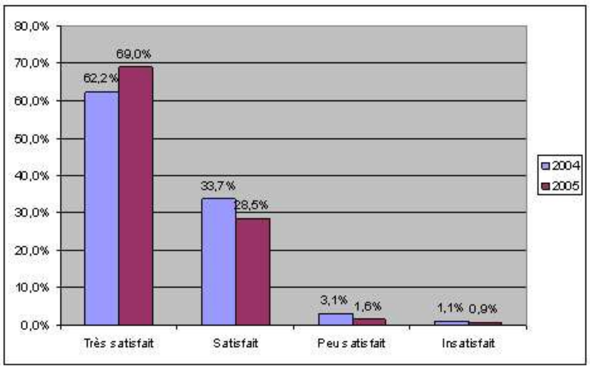
valeurs) et sa fréquence (ou son effectif), la hauteur du rectangle est proportionnelle à la fréquence (ou à l'effectif, ce qui revient au même). On pourrait finalement regrouper tous ces diagrammes sous le même terme générique de *diagramme en bâtons* en admettant toutes les apparences des bâtons : segments et rectangles contigus ou non. Les trois premiers graphiques analysés sont des exemples de ces *diagrammes en bâtons*.
Les diagrammes de la seconde catégorie, les histogrammes, sont aussi des diagrammes « à bandes rectangulaires », mais les rectangles sont nécessairement contigus. Dans un histogramme, l'axe des abscisses est l'axe des valeurs, il est impossible qu'il soit l'axe des individus. L'histogramme repose sur la proportionnalité entre la hauteur de la bande et la densité de fréquence associée à toutes les valeurs du caractère qui, réunies dans un intervalle, forment la largeur de la bande rectangulaire. En reprenant la terminologie des fonctions, un histogramme est la représentation graphique de la fonction qui à toute valeur du caractère associe sa densité de fréquence. Il en résulte, comme dans les exemples que nous avons vus, que le graphique obtenu est celui d'une fonction en escalier dont l'apparence est bien celle d'un diagramme à bandes rectangulaires. Jean-Claude Régnier (2005) propose que ne soient plus représentés les côtés des rectangles parallèles à l'axe des ordonnées, et que la représentation de l'histogramme devienne ainsi davantage conforme à sa définition : graphe d'une fonction en escalier. On retrouve la propriété qui est utilisée dans le cinquième graphique comme procédé de représentation : dans un histogramme, l'aire de chaque bande rectangulaire est proportionnelle à la fréquence (ou à l'effectif) de la classe qu'elle représente.
Pour conclure ce paragraphe, indiquons que ces définitions ne se retrouvent pas dans tous les ouvrages consultés, loin de là ! La définition de l'histogramme que nous présentons est conforme à celle qui se trouve dans les ouvrages de statistique rédigés par des statisticiens mathématiciens et dans certains ouvrages pour les sciences humaines. On la trouve aussi dans le dictionnaire *Le petit Robert* : « *Graphique représentant la densité d'un effectif en fonction des valeurs d'un caractère, et formé par une série de rectangles dont la base constitue un intervalle de variation de ces valeurs et la surface l'effectif correspondant*. » La définition où les classes ont la même longueur et où l'axe des ordonnées représente les fréquences de ces classes se trouve dans les ouvrages de vulgarisation ou dans les autres ouvrages pour les sciences humaines. C'est aussi celle qu'on trouve dans le dictionnaire *Larousse* : « *Représentation graphique des classes d'une variable statistique, associant à chaque classe un rectangle proportionnel par sa longueur à l'amplitude, par sa hauteur à l'effectif de cette classe*. » C'est encore celle qui est implicite dans les tableurs comme Excel. Enfin signalons que la définition qui présente l'histogramme comme un diagramme où les aires des rectangles sont proportionnelles aux effectifs ou aux fréquences se rencontre presque uniquement dans des manuels ou sur les sites relatifs à l'enseignement secondaire. Nous interprétons cette définition comme un produit de la transposition didactique, lorsque la densité n'est pas enseignée.
Le lecteur remarquera comme nous que les différentes définitions ne sont pas cohérentes mais que certaines ne sont pas non plus totalement contradictoires. Comment trancher ? Suivant la proposition de Gérard Vergnaud (1991) « *la connaissance rationnelle est opératoire ou n'est pas* », nous nous proposons, afin de répondre à la question, de poursuivre par l'étude de l'activité relative à ces graphiques.

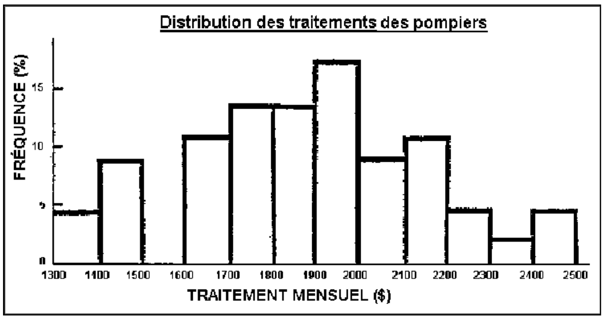
1.2. La question des activités relatives aux diagrammes rectangulaires

Les recherches menées en didactique des mathématiques sont plus nombreuses à aborder la question des activités relatives aux dessins et aux figures géométriques qu'elles ne le sont à traiter celle des activités relatives aux graphiques et aux diagrammes.
En nous inspirant à la fois des travaux de Raymond Duval (2005) qui décrit quatre niveaux d'activité sur les figures géométriques (botaniste, arpenteur, constructeur et inventeur-bricoleur) et des travaux de Dominique Lahanier-Reuter (2005) qui portent sur l'histogramme et des difficultés que rencontrent les élèves avec ce diagramme, nous proposons de distinguer deux types d'activités :

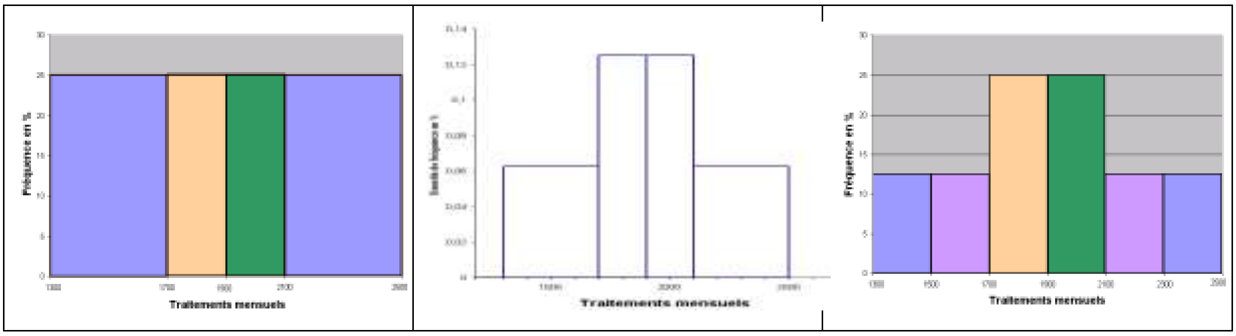
- les activités « iconiques » où le sujet reconnaît et/ou interprète le diagramme pour sa forme ;
- et les activités « graphiques » où le sujet réalise et/ou étudie le diagramme par des calculs, des mesures, des constructions ou des comparaisons.

Illustrons notre propos en reprenant les graphiques présentés précédemment.
Une activité de type iconique portant sur le diagramme représentant les niveaux de satisfaction des jeunes après leur séjour scientifique conduit à dégager à la fois le grand niveau de satisfaction des jeunes ainsi que la croissance de cette satisfaction de 2004 à 2005. Le diagramme des traitements mensuels des pompiers québécois, pour le sujet qui possède quelques connaissances de la loi normale, évoquera, dans le cadre d'une activité de type iconique, l'allure gaussienne de la répartition avec un salaire mensuel moyen et un écart-type approximativement égaux à respectivement 1 900 \$ et 300 \$. Ce type d'activité conduira aussi à remarquer l'absence de salaires compris entre 1 500 et 1 600 dollars canadiens.





Une activité de type graphique portant sur le même diagramme pourrait conduire à regrouper les salaires en quatre classes : [1 300 ; 1 700[, [1 700 ; 1 900[, [1 900 ; 2 100[et [2 100 ; 2 500[car la fréquence de chaque classe est approximativement égale à 25%. Si le graphique est un diagramme en bâtons (en tuyaux d'orgue), alors l'axe des ordonnées est l'axe des fréquences et le graphique obtenu est celui représenté ci-après à gauche. La lecture de ce graphique est ambiguë : l'égalité des ordonnées permet de lire l'égalité des fréquences mais d'une part cette égalité des fréquences est mal perçue tant les rectangles apparaissent comme différents, et d'autre part, il a perdu sa forme en cloche. Si le graphique est un histogramme, alors l'axe des ordonnées est l'axe des densités de fréquence et le graphique obtenu est celui du centre, qui rend mieux compte de l'égalité des fréquences grâce à la ressemblance des rectangles et qui conserve la forme en cloche du premier graphique. La différence entre ces deux graphiques disparaîtrait en partageant les deux classes extrêmes du premier en deux pour obtenir non plus quatre classes d'amplitudes inégales mais six classes de même amplitude, c'est ce que montre le graphique de droite.



En définitive, dans l'histogramme central, on pourrait aussi choisir de représenter la classe des salaires compris entre 1 700 \$ et 2 100 \$ et avoir ainsi trois classes de même amplitude 400 \$. La classe centrale représentant 50% des salaires et les deux classes latérales 25% chacune. Le lecteur comprendra ainsi pourquoi, seule la définition des histogrammes comme représentation de la densité de fréquence est compatible avec des activités de type graphique conduisant à regrouper ou à scinder des classes : réunir deux classes adjacentes d'un histogramme conduit à réunir les rectangles qui les représentent, cela ne change pas la forme de l'histogramme. Ce n'est pas le cas pour un diagramme en tuyaux d'orgue, sauf à imposer que les classes soient toutes de la même longueur c'est-à-dire à

empêcher en partie l'activité graphique sur ce diagramme. Une dernière remarque à ce sujet : la définition proposée dans certains manuels de l'enseignement secondaire (aires proportionnelles aux fréquences) convient également pour mener ces activités de type graphique, est en effet cohérente avec celle qui repose sur la densité de fréquence, mais elle évite (vraisemblablement à dessein) d'introduire cette notion.
Indiquons enfin que le choix du nombre de classes lors de la construction d'un histogramme relève bien d'une activité de type graphique. Comme l'expose Dominique Lahanier-Reuter dans le cours de statistique qu'elle donne à l'université Lille 3, le choix du nombre de classes n'est pas indépendant de l'interprétation qu'on veut produire de la distribution. Le graphique initial des salaires des pompiers québécois évoque une courbe gaussienne. Comme l'ont montré les graphiques précédents, en diminuant le nombre de classes à quatre, l'histogramme présente une répartition des salaires qui peut s'interpréter en évoquant les parts équivalentes des salaires faibles, plutôt faibles, plutôt élevés et élevés. Avec un nombre pair de classes, on est en effet conduit à opposer les salaires faibles et les salaires élevés. En choisissant un nombre impair de classes, trois classes par exemple, on est conduit à dégager une classe centrale autour de laquelle s'opposent les classes latérales : la moitié des pompiers touche un salaire moyennement élevé compris entre 1 700 $ et 2 100 $, un quart d'entre eux touche un salaire faible compris entre 1 300 $ et 1 700 $ et le dernier quart touche un salaire élevé compris entre 1 700 $ et 2 100 $.
Les recherches en didactique des mathématiques ont montré qu'il est indispensable que les élèves comprennent comment utiliser les dessins et les figures pour qu'ils puissent résoudre des problèmes géométriques. Et pour cela, l'enseignement doit leur proposer des activités sur ces dessins et ces figures. De même, nous pensons qu'il est nécessaire que les élèves développent des activités de type iconique et de type graphique sur les diagrammes afin qu'ils puissent les utiliser pour résoudre les problèmes qu'ils rencontreront dès qu'ils auront à analyser ou à communiquer des données. En ce qui concerne l'histogramme et son utilisation, la définition référant à la densité de fréquence permet les activités de type graphique que les élèves devraient développer, la confusion avec le diagramme en tuyaux d'orgue empêche au contraire ces activités.

2. Questions sur l'enseignement et sur l'apprentissage de l'histogramme

Quels sont les choix de l'institution scolaire quant à l'enseignement de l'histogramme ? Quels sont les types de tâches proposés aux élèves ? Comment les élèves les réussissent-ils ? Les réponses à ces interrogations montrent-elles que les élèves apprennent, au long de la scolarité, à interpréter, construire, et utiliser les histogrammes ? La seconde partie de ce texte aborde ces questions et propose un développement quant à un enseignement possible des histogrammes favorisant les deux types d'activités, iconiques et graphiques.

2.1. L'enseignement de l'histogramme

Commençons par identifier, à travers quelques citations des programmes et documents d'accompagnement, les choix actuels de l'institution scolaire française quant à l'enseignement de l'histogramme. Nous examinerons alors les propositions des auteurs de manuels scolaires et les choix globaux des enseignants du secondaire.

2.1.1 Quelques repères sur les programmes d'enseignement

Voici quelques extraits des programmes en cours en 2007-2008, les références sont indiquées en bibliographie.
L'enseignement des diagrammes à bandes est programmé dès la fin de l'école élémentaire :

> À l'école primaire, les élèves sont amenés à lire, interpréter et utiliser divers modes de représentations de données (liste, tableaux, diagrammes, graphiques). Ce travail se poursuit au collège dans le domaine des statistiques. A l'école, comme au collège, l'analyse critique des informations communiquées à travers de tels supports participe à la formation du citoyen.

Cet enseignement se poursuit donc au collège dans le cadre de l'enseignement de l'organisation des données et de la statistique.

> Lire et interpréter des informations à partir d'un tableau ou d'une représentation graphique (diagrammes divers, histogrammes). Présenter des données sous la forme d'un tableau, les représenter sous la forme d'un diagramme ou d'un histogramme. Le choix de la représentation est lié à la nature de la situation étudiée. Pour les données relatives à un caractère qualitatif trois types de représentations graphiques sont utilisés : le diagramme en tuyaux d'orgue, le diagramme en bandes (ou diagramme linéaire), le diagramme à secteurs (circulaires ou semi-circulaires). Pour les données à caractère quantitatif discret (ou à valeurs discontinues) le diagramme utilisé est le diagramme en bâtons ; pour les données à caractère continu, un histogramme est utilisé (en se limitant au cas de classes d'égale amplitude).

Puis au lycée, en classe de 1re dans la série scientifique et dans la série économique et sociale, le cas des histogrammes comportant des classes d'amplitudes différentes est abordé.

> Les histogrammes à pas non constants ne seront pas développés pour eux-mêmes, mais le regroupement en classes inégales s'imposera lors de l'étude d'exemples comme des pyramides des âges ou de salaires.

Pour l'ensemble des programmes, parallèlement à la transmission de ces savoirs en statistique, les professeurs sont explicitement chargés de l'éducation citoyenne de leurs élèves tant par les thèmes abordés dans les problèmes que par l'enseignement même de la lecture d'images.

2.1.2. Propositions des auteurs de manuels scolaires

Tenant compte de ces prescriptions, les auteurs des manuels scolaires du secondaire proposent à la fois des informations sur la notion d'histogramme et des exercices pour mettre en œuvre cette notion. Les manuels scolaires indiquent tous (au moins implicitement) que les histogrammes représentent la distribution d'une variable numérique continue, que les bandes rectangulaires qui le composent sont contiguës, ils peuvent néanmoins se distinguer en deux catégories suivant la définition proposée :

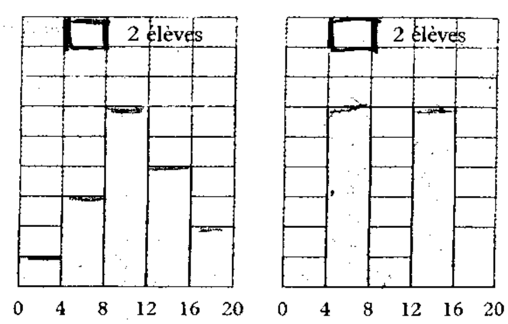
- soit la définition repose sur la proportionnalité des hauteurs des bandes rectangulaires aux effectifs ou aux fréquences des modalités, avec contrainte d'égalité des largeurs des bandes ;
- soit la définition repose sur la proportionnalité des aires des bandes aux effectifs ou aux fréquences, sans contrainte alors sur l'égalité des largeurs des bandes rectangulaires. Dans ce dernier cas, l'axe des ordonnées est absent et l'unité est représentée à l'aide d'une surface rectangulaire, comme sur le cinquième exemple étudié dans le paragraphe précédent. En revanche, les auteurs précisent que si les bandes rectangulaires ont la même largeur, alors les effectifs ou les fréquences sont aussi proportionnelles aux hauteurs des bandes.

Nous constatons cependant que ces choix ne sont pas justifiés, ni questionnés, et que certains auteurs pourraient même décourager d'avance la réflexion des lecteurs (élèves ou professeurs) quant à la signification de l'axe des ordonnées d'un histogramme dont les bandes rectangulaires sont de largeurs inégales (cf. Hachette, col. Terracher, 2nde, édition 1994).

> ■ L'histogramme (*cf.* exemple 3)
> Dans un histogramme, les effectifs (ou les fréquences) et les aires des rectangles sont proportionnels. Et donc, *lorsque les classes ne sont pas de même amplitude,* il est parfaitement « imbécile » de proposer une unité quelconque sur l'« axe des ordonnées ».

Les tâches proposées dans les exercices portent principalement sur le passage des données tabulaires aux données graphiques et vice-versa. Après le travail de traduction, les questions portent souvent sur la détermination de paramètres de position (cf. Didier, col.Math'*x*, 2nde, édition 2005 et Nathan, col. Transmath, 2nde, édition 2004).

2 On a représenté ci-dessous les résultats de deux classes au dernier contrôle de maths :



1. Donner le tableau des effectifs des deux séries de notes ainsi représentées.

2. L'une des série peut être qualifiée de « bimodale ». Laquelle et pourquoi ? À quoi cela correspond-il concrètement dans la classe ?

Classes et histogrammes

18 Voici les tailles en cm des 30 élèves d'une classe de Seconde : 1,62 ; 1,62 ; 1,73 ; 1,84 ; 1,56 ; 1,64 ; 1,69 ; 1,74 ; 1,74 ; 1,70 ; 1,66 ; 1,68 ; 1,71 ; 1,72 ; 1,79 ; 1,60 ; 1,62 ; 1,78 ; 1,79 ; 1,65 ; 1,65 ; 1,71 ; 1,74 ; 1,78 ; 1,84 ; 1,65 ; 1,69 ; 1,68 ; 1,68 ; 1,66.

1. Recopiez et complétez le tableau suivant, dans lequel toutes les classes ont une amplitude de 5 cm.

classe	[155 ; 160[	[160 ; 165[	...	[180 ; 185[
effectif	1	...	...	...

2. Quelle est la classe modale de cette série ainsi classée ?

3. Construisez l'histogramme de cette série classée.

2.1.3. L'histogramme dans les classes de collège et de lycée

Nous savons que l'écart peut être grand entre ce qui figure dans les manuels et ce qui est effectivement proposé en classe (dans un sens comme dans l'autre d'ailleurs). Bien que nous n'ayons pas enquêté de manière organisée auprès des professeurs, les avis exprimés indiquent que la notion n'est pas intéressante pour poser des problèmes mathématiques, mais qu'elle permet néanmoins de proposer des situations d'application de la proportionnalité. La durée de l'enseignement de la notion est très courte, souvent moins d'une heure.

Les professeurs savent que quelque chose de difficile est en jeu dans la définition de l'histogramme à cause du cas où les largeurs des bandes rectangulaires ne sont pas de longueurs égales, mais la connaissance de la notion de densité de fréquence n'est pas connue par tous les professeurs de l'enseignement secondaire, même ceux qui enseignent les probabilités au lycée.

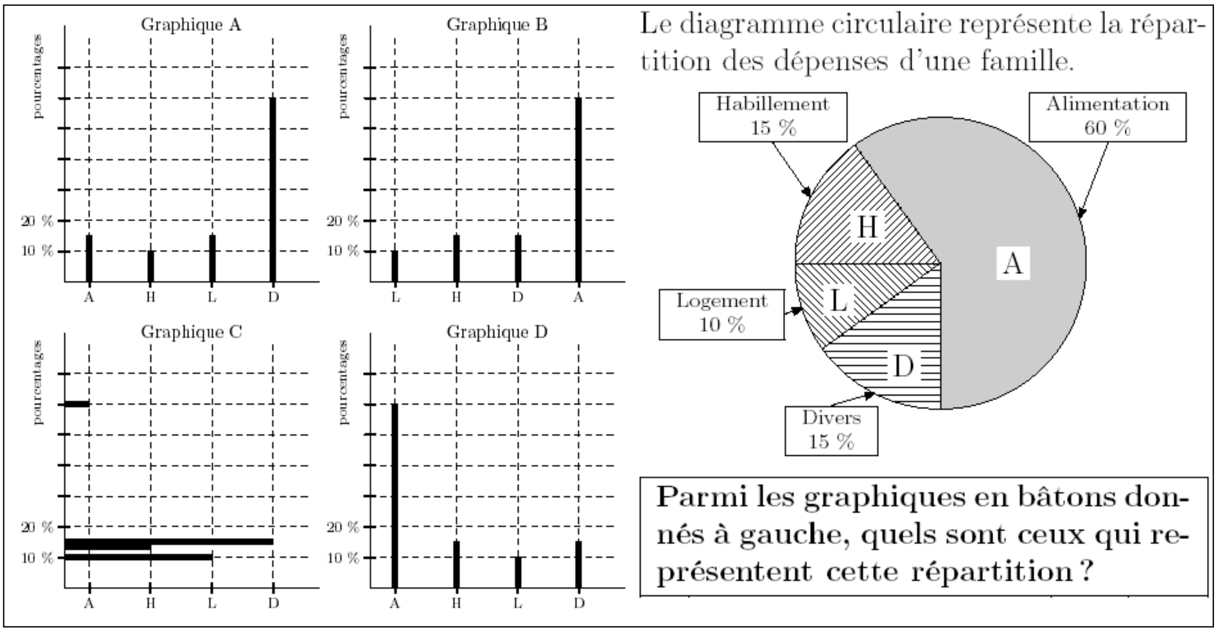
Citons comme témoignage du manque d'intérêt pour la notion, l'article de Claudine Schwartz & Jacques Treinerla (2006) consacré à la lecture de l'histogramme publié par le site « *Statistix* » dédié à l'enseignement de la statistique au secondaire et animé par une équipe de l'INRP proche de l'institution scolaire. Le titre même de la page est évocateur : « Que dire d'un histogramme ? » Tout comme le texte situé sous le lien vers l'article : « *Cette question est souvent posée par des professeurs en charge d'un enseignement de statistique descriptive.* » L'article propose une maïeutique entre le narrateur S et un maître M qui, ayant obtenu un histogramme après collecte de données auprès de ses étudiants, ne sait quoi en dire. S parvient à ce que M change de demande, il lui propose entre autres de ne pas chercher à interpréter un histogramme, mais plutôt à en comparer plusieurs…

2.2. La question de l'apprentissage

Nous n'avons pas mené d'enquête systématique sur l'apprentissage de la notion d'histogramme car l'étude préliminaire a montré que nous pouvions nous attendre à ce que les connaissances des élèves soient limitées à ce qui est principalement enseigné c'est-à-dire les traductions entre données tabulaires et des données graphiques. Deux questions posées dans les évaluations menées par l'APMEP (disponibles en ligne sur le site EVAPMIB référencé en bibliographie) montrent que les élèves ne sont pas familiers des activités sur les histogrammes. Une analyse de copies d'étudiants en 2e année de Licence de mathématiques appliquées aux sciences sociales (MASS) montre que beaucoup d'étudiants possèdent mal cette notion.

2.2.1. Confusion entre tâche iconique et tâche graphique

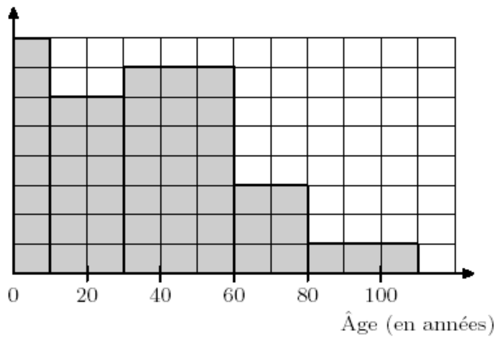
Bien qu'elle ne concerne pas directement les histogrammes, une question posée en 1990 aux élèves de la classe de 5e lors d'une évaluation menée par l'APMEP a retenu notre attention car elle montre que la capacité à interpréter des graphiques nécessite un enseignement pour être acquise (référence permettant d'accéder à la question sur le site : clé = 482). Il s'agissait de reconnaître les diagrammes en bâtons correspondant à un diagramme circulaire.



La réponse correcte, B et D, est proposée par seulement 45% des élèves, la plupart de ceux qui échouent proposent la réponse A et D. Comment interpréter cette réponse trop fréquente pour être considérée comme accidentelle ? Les professeurs de l'association proposent une explication peu convaincante : les élèves auraient choisi les graphiques symétriques. En interrogeant leurs élèves à ce sujet, les professeurs du groupe didactique de l'IREM de Paris ont obtenu une interprétation qui montre mieux comment la tâche est réalisée par ceux qui proposent cette réponse A et D. Dans le graphique circulaire, on observe une très grande part ainsi qu'une petite part comprise entre deux parts plus grandes. Sur les graphiques A et D, la situation est la même : un très grand bâton, et à côté, un petit bâton compris entre deux plus grands. Autrement dit les élèves auraient recherché à reconnaître la forme du diagramme circulaire dans le diagramme en bâtons. Ils ont développé une activité iconique et non une activité graphique. Une telle activité aurait permis, par exemple, de remarquer que dans le graphique A ce n'est pas la part « Logement » qui est la plus petite, alors que c'est le cas dans le diagramme circulaire et dans le graphique D. De nombreux élèves n'ont donc pas tenu compte, pour réaliser cette tâche, des informations portées sur les axes ou en légende, et ont retenu seulement des caractéristiques liées à la forme des graphiques proposés.

2.2.2. L'absence d'axe des ordonnées rend la lecture difficile

Une question (référence permettant d'accéder à la question sur le site : clé = 713) de lecture graphique a été posée en 1991 à des élèves en classe de 2nde, il s'agissait « simplement » de déterminer la fréquence d'une classe correspondant à une bande sur un histogramme ne comportant pas d'information sur l'axe des ordonnées.



Cet histogramme représente la répartition des âges des habitants d'un quartier.

Quel est le pourcentage des habitants de ce quartier dont l'âge est compris entre 30 et 60 ans ?

Réponse :

....

Une analyse rapide de la tâche permet de repérer que les élèves doivent d'abord déterminer qu'un carreau représente 2% de la population en constatant que la population est entièrement représentée par les 50 carreaux du graphique. Cela leur permet de trouver que la fréquence de la classe [30 ; 60[qui en comporte 21 est égale à 42%. L'absence d'indication quant à la valeur d'un carreau a sûrement déstabilisé les élèves à qui l'histogramme a été présenté seulement avec la contrainte d'égalité des largeurs des bandes rectangulaires et avec un axe des ordonnées qui indique les effectifs ou les fréquences. Ainsi seulement 31% des élèves déterminent le pourcentage demandé.

De nombreux élèves échouent car ils tiennent compte de la hauteur seulement (ils comptent 7 carreaux et annoncent une fréquence de 70%) ou de la largeur (ils comptent 3 carreaux et annoncent une fréquence de 30%). Certains attribuent la valeur 100% soit à la hauteur maximale (ils calculent alors 7/8 de 100% et annoncent une fréquence de 87,5%) soit à la largeur totale (ils calculent alors 3/11 de 100% et annoncent une fréquence de 27%). Enfin certains ne tiennent pas compte de la longueur des classes et associent la valeur 100% au cumul des hauteurs des bandes rectangulaires (ils obtiennent alors une hauteur totale de 8+6+7+3+1=25, ils calculent 7/25 de 100% et annoncent une fréquence de 28%).

Les procédures des élèves montrent d'une part qu'ils ne manquent pas de ressources en calcul pour aborder ces questions, ils seraient donc vraisemblablement capables d'utiliser une information sur l'axe des ordonnées comme « Densité de fréquence : % pour 10 ans » avec les valeurs 2, 4, 6, 8, etc. à chaque graduation. Mais encore faudrait-il que la notion de densité de fréquence soit enseignée.

Nous retrouvons ce que mentionnait Dominique Lahanier-Reuter (2005) que nous avons déjà citée et que nous interprétons de la façon suivante : l'absence d'indication de la densité de fréquence sur l'axe des ordonnées nuit à lecture d'un histogramme.

2.2.3. Confusions à propos de l'histogramme en L2 MASS

L'enseignement secondaire de l'histogramme porte essentiellement sur la traduction entre diagramme et tableau de valeurs. Les manuels proposent essentiellement ces tâches. Les professeurs déclarent les donner à leurs élèves pour mettre en œuvre la proportionnalité et pour répondre à la demande ministérielle d'éducation citoyenne. Ces éléments laissent penser que les élèves qui ont suivi sans échec l'enseignement de mathématiques du secondaire savent au moins construire un histogramme à partir d'un tableau et inversement. Néanmoins, la durée d'enseignement que les professeurs déclarent semble tellement limitée, que nous avons souhaité savoir si les élèves sortis « par la grande porte » du système scolaire savent effectivement réaliser un tel diagramme et, pour ceux qui n'y parviennent pas, connaître leurs difficultés.

Un enseignant chercheur, chargé du cours de statistique en L2 de mathématiques appliquées aux sciences sociales et qui déclare avoir le souci de l'enseignement de la statistique descriptive, a proposé une question portant sur l'histogramme dans ses sujets de devoir et d'examen. Les deux fois, un tableau avec de nombreuses données était fourni, un résumé graphique par un histogramme était demandé. Nous avons étudié les 79 copies de ces étudiants. Nous avons accepté les représentations correspondant aux deux définitions rencontrées dans l'enseignement secondaire : représentation de la distribution des fréquences ou des effectifs par un diagramme à bandes rectangulaires contiguës de même largeur et dont les hauteurs sont proportionnelles aux fréquences ou aux effectifs, ou bien par

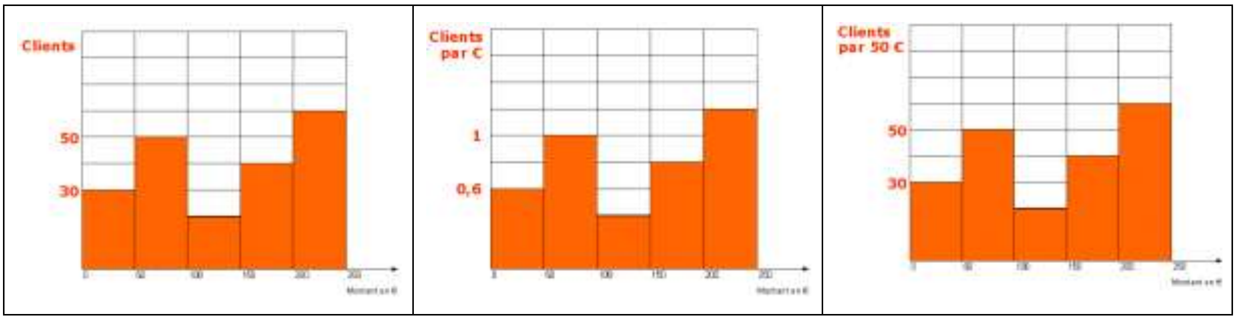
un diagramme à bandes rectangulaires contiguës dont les aires sont proportionnelles aux fréquences ou aux effectifs.
Nous avons obtenu seulement une moitié de réponses correctes (42 soit 53%) dont une seule propose un histogramme où ni les fréquences, ni les effectifs ne figurent sur l'axe des ordonnées : une unité d'aire est indiquée sur le graphique. Aucune des réponses ne représente la densité des fréquences ou des effectifs de la variable. Aucune réponse ne représente des classes d'amplitudes différentes. Certains étudiants (7 soit 9%) proposent un diagramme correct quant à la hauteur des bandes rectangulaires, mais elles ne sont pas contiguës, l'axe des abscisses n'est pas gradué, les intervalles de valeurs correspondant à chaque bande sont indiqués sous la bande. D'autres étudiants (8 soit 10%) représentent la série statistique et indiquent par conséquent chaque individu et sa valeur par une étroite bande rectangulaire, les individus étant rangés par ordre d'apparition dans le tableau de valeurs. Enfin, parmi les autres étudiants (22 soit 28%), quelques-uns proposent un diagramme en bâtons et la plupart d'entre eux s'abstiennent de répondre alors que la question compte pour leur évaluation.
Ces résultats ont été jugés très décevants par l'enseignant chercheur qui pensait que les contenus de l'enseignement secondaire étaient acquis par davantage d'étudiants. Une réflexion s'est engagée quant à un éventuel approfondissement de son cours de statistique descriptive en MASS. Ces résultats, sans doute bien supérieurs à ceux que nous aurions obtenus sur un échantillon d'étudiants moins intéressés par les mathématiques, nous montrent que les objectifs de lecture et de réalisation de graphiques ne sont pas atteints à l'issue de l'enseignement secondaire.

3. Pour un enseignement des histogrammes au collège et au lycée

Dans ce dernier paragraphe, nous souhaitons tirer un bilan de l'étude menée et en discuter les résultats pour proposer des tâches mathématiques pour l'enseignement des histogrammes réparti entre le collège et le lycée, des tâches qui permettent aux élèves de s'engager dans des activités iconiques et graphiques.

3.1. L'histogramme, un objet à définir ?

Après avoir montré la diversité des graphiques qui sont, au sens courant, regroupés sous le terme d'histogramme, nous avons développé une analyse de ceux qui sont proposés pour l'enseignement secondaire. Deux définitions divergentes émergent. Selon les deux définitions, les histogrammes sont des diagrammes à bandes rectangulaires contiguës. En revanche, selon la première définition, l'histogramme représente la distribution des fréquences (ou des effectifs) d'une série statistique dont le caractère est numérique et continu ; l'axe des abscisses est l'axe des valeurs du caractère et l'axe des ordonnées celui des fréquences (des effectifs). Selon la seconde définition, l'histogramme représente la densité de fréquence d'une variable statistique numérique et continue.
Afin de mieux comprendre les conséquences de ces définitions, nous avons envisagé l'activité d'un sujet qui lit, construit, analyse ou interprète un histogramme. Nous avons distingué les activités « iconiques » où le sujet reconnaît et/ou interprète le diagramme pour sa forme des activités « graphiques » où le sujet réalise et/ou étudie le diagramme par des calculs, des mesures, des constructions ou des comparaisons.
Le bilan auquel nous parvenons est que la première définition permet une activité iconique à la condition que les classes soient de même amplitude, en revanche, une activité graphique n'est réellement possible qu'avec la seconde définition. D'autre part, nous avons indiqué qu'un histogramme réalisé avec la première définition et des classes de même amplitude pouvait s'interpréter facilement avec la seconde définition. Développons cette affirmation en reprenant le cinquième exemple présenté dans la première partie de l'article, l'histogramme représente la distribution des effectifs des dépenses effectuées sur un site Internet par un échantillon de 200 consommateurs. Le diagramme de gauche où l'axe des ordonnées indique des effectifs peut s'interpréter comme l'histogramme proposé au centre où l'axe des ordonnées indique une densité d'effectif (analogue à ce qu'on obtient de manière automatique avec un logiciel de statistique), ou plus simplement comme celui proposé à droite avec une autre unité pour cette densité.



Comme on sait que le nombre de clients qui constituent l'échantillon est 200, on peut aussi représenter la densité de fréquence de la variable plutôt que sa densité d'effectif ; la valeur correspondant à une graduation est alors 0,1 % par € c'est-à-dire 5% pour 50 euros. Dans le graphique de droite, la légende serait remplacée par « % pour 50 € » le nombre 30 serait remplacé par 15 et le nombre 50 par 25.
Pour en terminer sur la définition de l'histogramme, il nous semble qu'un compromis acceptable est de ne pas définir cet objet en fonction de ce qu'il représente mais en fonction de son procédé de représentation. Ainsi, nous appelons histogramme tout diagramme à bandes rectangulaires contiguës qui est caractérisé par deux axes de coordonnées. Nous réservons les qualificatifs « en bâtons » et « en tuyaux d'orgue » aux diagrammes pour lesquels les bandes rectangulaires ne sont pas contiguës ; ces deux diagrammes se distinguant seulement par la largeur des bandes rectangulaires, qui est nulle pour le diagramme en bâtons et qui ne l'est pas pour le diagramme en tuyaux d'orgue. Suivant ce qu'on veut représenter, le choix de tel ou tel graphique est adapté. On choisit un diagramme en bâtons ou en tuyaux d'orgue pour représenter une série statistique ou pour représenter la distribution d'une variable discrète. Les valeurs d'une variable continue étant regroupées en classes, on choisit un histogramme pour représenter sa densité ou sa distribution. Dans le dernier cas, afin d'éviter toute ambiguïté graphique, on choisit des classes de même amplitude, et à défaut d'en disposer, on préfèrera représenter sa densité.

3.2. Propositions de tâches pour le collège et pour le lycée

Les activités graphiques, comme nous l'avons montré, reposent sur la notion de densité. Les grandeurs quotients étant difficiles à acquérir et la notion de densité étant très abstraite, il ne nous semble pas raisonnable de l'envisager au collège. Ainsi, nous envisageons que l'enseignement propose des tâches conduisant à des activités iconiques au collège et à des activités des deux types au lycée.

3.2.1. Des tâches d'interprétation et de critique dès le collège

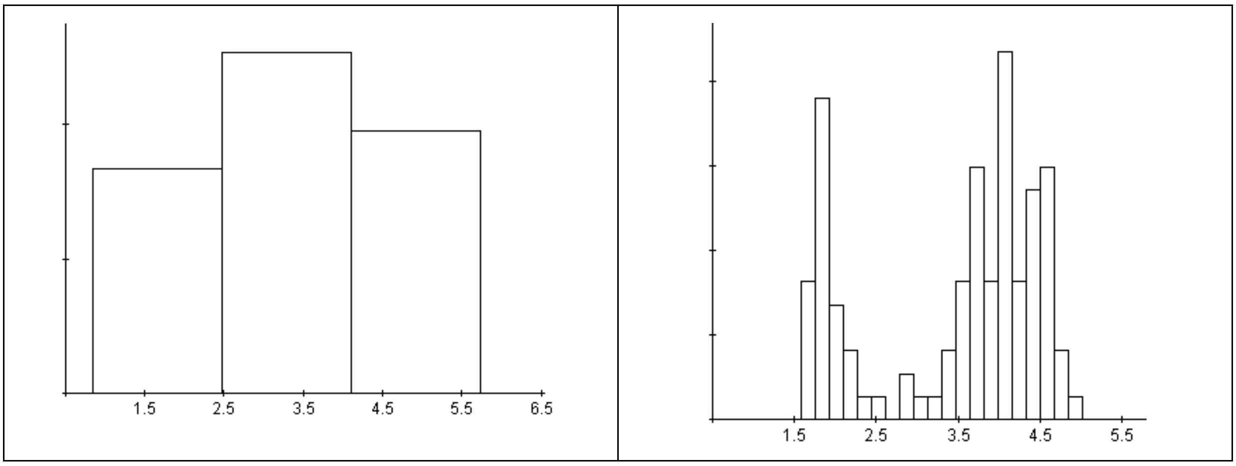
Dès le collège, nous envisageons un enseignement qui conduit les élèves à la réalisation comme à l'interprétation des histogrammes, qui pose la question de la fidélité du graphique, par exemple, en interrogeant le choix des classes. Les élèves apprendront comment lire et interpréter un graphique :

- identifier les grandeurs indiquées sur les axes et les unités ;
- repérer et attribuer une signification aux zones hautes, basses et étendues horizontalement, puis aux variations des hauteurs ;
- examiner la symétrie ou l'asymétrie de la forme globale constituée par l'ensemble des rectangles ;
- disposer d'histogrammes de référence pour interpréter un histogramme d'après sa forme globale par comparaison à des phénomènes déjà identifiés (distribution sur un échantillon d'une population où les valeurs du caractère sont équiréparties ou distribuées normalement, distribution sur mélange de deux échantillons issus de deux populations de paramètres différents, etc.)

Lorsqu'ils réalisent un histogramme, les données non classées pourront leur être confiées pour qu'ils apprennent à choisir le nombre de classes et leur amplitude. La situation suivante montre l'intérêt d'un tel travail.

> Dans un collège, il y a possibilité pour les élèves en difficulté d'assister librement à des études dirigées. Les progrès des élèves ont été relevés par le nombre de points d'augmentation de leur moyenne générale.

La même série statistique conduit en effet à ces deux histogrammes.



Le premier graphique montre que le dispositif a fait progresser les notes des élèves de manière variable, le plus fréquemment de 2,5 à 4 points. Le second graphique laisse apparaître deux groupes qui progressent différemment, l'un de 2 points environ et l'autre de 4 points environ. Les élèves assistant librement au dispositif de soutien, on peut se demander si les deux groupes ne sont pas précisément les assidus et les autres qui assisteraient à ces séances seulement à l'approche des évaluations…

3.2.2 Introduction à la notion de densité au lycée

Au lycée, nous envisageons de poursuivre par l'enseignement des histogrammes à partir de problèmes conduisant toujours à la réalisation d'histogrammes, à leur interprétation et à leur comparaison, ainsi qu'à la réflexion sur le choix des classes en fonction de la fidélité du résumé graphique.
Nous proposons également d'étendre ces activités aux mesures, aux calculs et aux constructions. En particulier, comme nous l'avons montré par l'étude des cinq exemples dans la première partie de ce texte, les situations conduisant à regrouper et à scinder des classes permettent d'engager une réflexion sur la stabilité du graphique. En référence à la situation, cette réflexion conduit à envisager la proportionnalité de la fréquence d'une classe à l'aire de la bande qui la représente et pas à sa hauteur. Enfin, l'interrogation sur l'axe des ordonnées conduit à la notion de densité de fréquence qui pourra servir de support à l'introduction des densités de probabilité le moment venu.

Pour conclure, indiquons que des expérimentations de scénarios d'enseignement sont en cours de réalisation pour évaluer la portée de nos propositions. Notre façon d'aborder cette question de l'enseignement et de l'apprentissage de l'histogramme repose sur une analyse mathématique approfondie du savoir à transmettre, mais pas seulement. Elle repose aussi sur une analyse conjointe du savoir et de son utilisation dans des problèmes accessibles aux élèves au niveau où le savoir est enseigné. Elle repose enfin sur une prise en compte des contraintes du métier de professeur de mathématiques. Il est en effet difficile d'enseigner une notion très isolée dans l'ensemble des programmes et pour laquelle les types de tâches proposés conduisent seulement à des activités de traduction. Nous proposons de relier cet enseignement à celui plus général des représentations graphiques dans l'enseignement secondaire, en indiquant des types de tâches propices au développement d'activités iconiques et graphiques variées : lecture, interprétation, critique, mesure, calcul, comparaison, synthèse, construction, etc.

Bibliographie

DUVAL R. (2005), Les conditions cognitives de l'apprentissage de la géométrie : développement de la visualisation, différenciation des raisonnements et coordination de leurs fonctionnements, *Annales de didactique et de sciences cognitives*, p. 5-53.

EVAPMIB, une base de questions d'évaluation en mathématiques.
http://ctug48.univ-fcomte.fr/evapmib/siteEvapmib/

LAHANIER-REUTER D. (2005), Quelles informations porte l'axe des ordonnées, *Statistiquement Vôtre*, n°6.
http://www.sfds.asso.fr/groupes/statvotre/StatVotre_6_2005.htm

REGNIER J.-C. (2005), Histogramme : Réflexion sur une représentation graphique particulière parfois abusivement utilisée tant dans l'enseignement que dans l'application de la statistique. *Statistiquement Vôtre*, n°6.
http://www.sfds.asso.fr/groupes/statvotre/StatVotre_6_2005.htm

SCHWARTZ C. & TREINER J. (2006), Que dire d'un histogramme ?, *in STATISTIX*.
http://www.statistix.fr/spip/spip.php?article18

VERGNAUD G. (1991), La théorie des champs conceptuels, *Recherches en didactique des mathématiques*, 10/2.3, p.133-170.

Liens vers les programmes mentionnés dans l'article

Programme du cycle 3 de l'enseignement primaire
ftp://trf.education.gouv.fr/pub/edutel/bo/2007/hs5/hs5_approfondissement.pdf

Programme du cycle central de l'enseignement secondaire
http://www.cndp.fr/produits/detailsimp.asp?Ref=755A2821

Documents d'accompagnement des programmes de mathématiques au collège : organisation et gestion de données (2007)
http://eduscol.education.fr/D0015/doc_acc_clg_organisation_donnees.pdf

Programmes des classes de première de l'enseignement secondaire
http://www.cndp.fr/archivage/valid/73275/73275-13626-17261.pdf