LA MOYENNE : UN CONCEPT INEXPLOITÉ, D'UNE RICHESSE EXCEPTIONNELLE

Linda GATTUSO
Département de mathématiques,
Université du Québec à Montréal

Il y a certains concepts qui, s'infiltrant doucement, deviennent tellement banals qu'on en parle à peine. La moyenne est un de ceux-là. Celle-ci, utilisée de façon extensive en statistique et dans les autres sciences, fait dorénavant partie de la vie quotidienne. La moyenne est un concept qui s'est développé assez récemment et est devenu essentiel au citoyen ordinaire comme au professionnel. Des informations statistiques nous arrivent quotidiennement sous forme de tableau, de graphiques ou d'énoncés qui relèvent de sondages ou d'enquête. Pour le technicien, le syndicaliste, le professionnel de la santé, le journaliste sans oublier l'universitaire, le chercheur, le scientifique, il est plus souvent nécessaire non pas de produire des statistiques mais bien de prendre des décisions en fonction des résultats obtenus.

Or, la moyenne est une caractéristique statistique très souvent mentionnée dans les articles scientifiques soit parce que l'on a recourt aux statistiques inférentielles ou à l'analyse des données (E.D.A.). Dans plusieurs contextes, elle est calculée une fois les données pondérées. Dans la prise de décisions, les méthodes d'analyse de coûts/bénéfices et de programmation linéaire l'utilisent également. Ce concept mérite certes que l'on s'y attarde.

Pourtant, une étude récente (Gattuso et Mary, 1996, 1997) montre que c'est un concept plutôt mal compris même chez des étudiants en formation des maîtres. Dans l'enseignement, la moyenne est trop souvent ramenée à une simple formule. Une petite somme, une petite division, et le tour est joué ! Par conséquent, la présentation traditionnelle de la moyenne relève plus d'une description de l'algorithme que d'une tentative d'amener l'élève à se construire un "sens" de la moyenne. "Il faut additionner toutes les données et diviser la somme par le nombre de données" se contente-t-on pratiquement de dire. Or, les erreurs observées résultent souvent d'une utilisation inadéquate d'une combinaison "somme-division". Dans les lignes qui suivent, nous allons présenter nos réflexions sur la moyenne, appuyées sur certaines observations que nous avons pu faire autant dans nos classes que lors d'une étude menée auprès d'élèves du secondaire à l'université (*id.*).

Au Québec, la moyenne arithmétique est abordée dès le primaire. Toutefois, ce n'est que dans la troisième année du secondaire (14 ans) que l'on étudie les mesures de tendance centrale, et en particulier, la moyenne de données groupées et pondérées. Pour situer le lecteur français, disons que les petits québécois sont à l'école primaire de 6 à 11 ans, de la première à la sixième année. Ils amorcent ensuite le cours secondaire qui comporte 5 années (12-16 ans). Certains d'entre eux poursuivent au cégep (collège d'enseignement général et professionnel) pour 2 ou 3 ans selon les programmes avant de s'inscrire, s'il y a lieu, à l'université.

Ceci dit, la moyenne est un concept beaucoup plus riche qu'on le croit de prime abord et nous allons tenter de vous en convaincre. Ce concept, comme beaucoup d'autres dans le domaine de la statistique, fait appel à des notions mathématiques qui comportent leurs difficultés propres. Néanmoins, certaines sont plus spécifiques à la moyenne qui a certaines propriétés qui lui sont particulières. Par conséquent, nous croyons que travailler la moyenne dépasse largement l'application simple d'un algorithme. Ce concept prend plusieurs sens. Il faut l'aborder dans des situations variées de façon à tirer profit de sa richesse. Nous illustrerons nos propos par quelques exemples pouvant être exploités en classe.

Divers éléments en jeu dans la moyenne

Il reste que certaines des difficultés rencontrées dans les problèmes de moyenne ne sont pas spécifiques à ce concept. Comme on le sait, pour trouver une moyenne arithmétique, on peut faire la somme des données et ensuite diviser le résultat par le nombre de données. Cependant, nous ne reprendrons pas ici toutes les difficultés mathématiques inhérentes aux opérations d'addition et de division (Brousseau, G., 1988 ; Vergnaud, G., 1990). Nous ne reprendrons pas non plus les études déjà bien connues qui ont exploré les nombres décimaux (Bolon, J., 1992 ; Perrin, M-J. 1986 ; Brousseau, G., 1981) ou encore les structures multiplicatives (Vergnaud, G. Ricco, G. Rouchier, A. *et al.*, 1978). Ces éléments sont tous présents dans les calculs de moyenne. Nous savons aussi que pour certains le problème sera plus difficile si les données sont décimales ou tout simplement de grands nombres. Les fréquences et parfois même les données étant souvent exprimées en pourcentage ou en fraction une bonne compréhension des pourcentages et des rationnels est aussi nécessaire. Les problèmes de moyenne peuvent également faire appel à un rapport. Ne citons que cet exemple : "La moyenne d'âge d'un groupe de médecins et d'avocats est 40 ans. Si l'âge moyen des médecins est 35 et celui des avocats 50, quel est le *rapport* entre le nombre de médecins et le nombre d'avocats ?" (MAA Problem book IV, p. 3, n° 12, 1973).

En plus, comme dans tous les cas de problèmes à texte, une bonne compréhension de l'énoncé du problème est essentielle. Cependant, pour la moyenne cela se complique à cause de la variété des formulations utilisées. On parle de la "moyenne d'âge *par enfant*"" du "nombre moyen de clients *par jour* ", mais, on dit aussi : "le poids moyen des personnes..." au lieu de "poids moyen *par* personne..." ou encore tout simplement "quel est l'âge moyen ?".

Et encore dans le cas de problèmes de moyenne, on ne peut se limiter à la lecture et la compréhension du texte de l'énoncé, il faut aussi être familier avec les tableaux et les diverses représentations graphiques avec lesquels on présente les données. Un élève qui voit le problème suivant :

"La fréquence de biscuits brisés dans un paquet a été observée et on a accumulé les données suivantes :

Nombre de biscuits brisés	Fréquence
2	5
4	8
6	7
9	2
13	1

Combien y a-t-il en moyenne de biscuits brisés par paquet ?" (Gattuso et Mary, 1997)

peut, s'il sait bien lire le tableau, réussir à le résoudre même s'il ne connaît pas la formule $\frac{\Sigma x_i \times n_i}{n}$ où x_i représente la valeur de chaque donnée, n_i les effectifs de chacune des données et n le nombre total d'effectifs. Nous avons vu certains élèves procéder à une énumération et trouver ensuite la moyenne :

« $\frac{1+2+3+3+4+4+\ldots+20+20+20+20}{76}$ » (Gattuso *et al., id.*).

En fin de compte, les contextes choisis ne sont pas sans influence. Certains mieux connus des élèves peuvent faciliter la compréhension du problème et donner lieu à un plus haut taux de réussite. Dans le cas des problèmes de moyenne, le type de données (liées au contexte) joue aussi. Les problèmes où l'on traite du poids des personnes sont mieux réussis, par exemple, que ceux où les données sont une cote pour la note de chaque session suivie au collège (Pollastek, 1981). Cela pourrait s'expliquer par le fait que les personnes sont des entités concrètes et qui se visualisent mieux que les sessions scolaires.

Mentionnons également l'algèbre et toutes les difficultés et erreurs qui y sont liées. De très nombreux problèmes de moyenne peuvent être résolus algébriquement, mais nous verrons qu'il est parfois plus simple de procéder arithmétiquement ou encore graphiquement. Force est de constater que travailler la moyenne peut donner l'occasion d'activer plusieurs notions mathématiques.

Les difficultés spécifiques au concept lui-même

Jusqu'à présent, nous n'avons fait que soulever des difficultés qui, pour la plupart, se retrouvent ailleurs que dans des problèmes de moyenne. Il reste que la moyenne a ses propres écueils. Un des premiers tient au fait que tout le monde a une idée de ce qu'est une moyenne (1) et que le mot lui-même "moyenne" est un terme utilisé couramment. C'est ce qui porte à confusion car on ne lui donne pas toujours un sens précis. Même les dictionnaires ne

(1) Il ne sera question ici que de moyenne arithmétique

présentent pas toujours des définitions justes de ce concept mathématique. On retrouve par exemple dans le *Petit Larousse illustré 1988* : "Quantité qui tient le milieu entre plusieurs autres" et dans le *Petit Robert* : "Type également éloigné des extrêmes, généralement le plus courant". "Être dans la bonne moyenne". "Vous avez autrement de poigne que la moyenne des patrons". Il n'est donc pas étonnant que certains enfants tiennent au point de départ des conceptions inadéquates de la moyenne. Heureusement, les définitions citées en côtoient d'autres qui sont plus justes : "Nombre exprimant la valeur qu'aurait chacune des parties d'une somme si, la somme restant la même, toutes les parties étaient égales entre elles" (Larousse, 1988) et "En moyenne, en compensant les unes par les autres les différences en sens opposé" (*Petit Larousse illustré*, 1988).

Les conceptions diverses que l'on a observées chez les enfants et parfois même chez les adultes (Mokros, Russell, 1991, 1995) sont inadéquates dans certaines situations mais justes ailleurs. Par exemple, certains sujets croient que la moyenne est la donnée du milieu (médiane). Cependant, il arrive aussi que la moyenne soit la donnée du milieu comme dans ce cas : 4, 4, 5, 7, 9, 10, 10, où 7 est à la fois la moyenne et la donnée du milieu. D'autres pensent que la donnée qui apparaît le plus souvent (mode) est la moyenne, c'est-à-dire, ce qui arrive la "plupart du temps". Encore ici, c'est possible, si les données sont 1, 4, 4, 4, 4, 4, 4, 7. La moyenne est 4 et six des huit données valent 4. Les enfants voient aussi la moyenne comme une valeur plausible et lui donne le sens de "en général". Ils s'attendent donc à ce que la moyenne soit une donnée possible ou encore, une des données de la liste présentée. Raison de plus pour ne pas accepter qu'il y ait en moyenne "2,3 enfants par famille" ou encore que la moyenne de 1, 2, 4, 5 soit 3 qui n'est pas une des données.

Non seulement, il faut accepter que la moyenne puisse ne pas faire partie des données, mais aussi qu'elle puisse être fractionnaire et ne correspondre à aucune réalité comme 3,4 bonbons (par enfant) ou 2,6 enfants par famille ! Or, les résultats de notre étude (Gattuso et Mary, 1996, 1997) ont révélé que 23,5 % des élèves de secondaire 3 (14 ans), 10 % des collégiens (17-19 ans) et 22 % des étudiants de l'université (19^+ ans) évitent de donner une moyenne fractionnaire [2]. Dans ce cas [3], il s'agissait d'un nombre de bonbons qui dans la réalité n'est habituellement pas fractionnaire. Il est rare, en effet, que l'on parle d'un quart de bonbon ! Ces réponses étaient accompagnées des commentaires suivants : "8,25 bonbons veut dire une moyenne de 8 bonbons par enfant et il en reste un..." ou encore "Trois des enfants ont 8 bonbons et le quatrième en a 9..."

Il est donc important de prime abord de sonder les conceptions des élèves et de choisir attentivement les situations d'apprentissage, en ayant en vue autant d'élargir le concept lui-même que de débusquer les conceptions erronées.

(2) Au Québec, la troisième année du secondaire est équivalente à la 4e française et le collège est équivalent au lycée.

(3) Lucie, Jacques et Paul se réunissent pour une fête. Ils ont apporté en moyenne 11 bonbons par enfants. Un quatrième enfant est arrivé mais il n'a pas apporté de bonbons. Quelle est maintenant la moyenne de bonbons par enfant ?

Certaines propriétés de la moyenne

Une bonne compréhension de la moyenne suppose également que l'on comprenne les propriétés de celle-ci. Une première propriété veut que la moyenne se situe entre la plus petite et la plus grande des valeurs de la distribution. Or, il n'est pas rare de voir nos élèves trouver des moyennes supérieures à la plus grande donnée, particulièrement lorsque les données sont regroupées et présentées dans un tableau de distribution avec de gros effectifs (4). Une deuxième propriété très importante dans certaines situations veut que la somme des écarts de chaque donnée à la moyenne soit nulle, $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$. Une autre propriété dit que la moyenne est la valeur dont la distance à toutes les autres données est minimale. En cela, elle est représentative de toutes les données et elle est influencée par toutes les données et toute modification sur l'une ou plusieurs des données influence la moyenne. Ceci est particulièrement à souligner dans le cas où l'une des données est nulle (5). Ce n'est pas parce que la somme reste inchangée quand l'on ajoute ou retranche 0 que la moyenne elle ne varie pas. Certaines études menées auprès d'enfants de 8 à 14 ans montrent que ces propriétés sont plutôt difficiles à saisir (Strauss, Bichler, 1988 ; Leon, Zawojewski, 1990).

Alors que certains élèves ne comprennent pas bien les propriétés de la moyenne, d'autres lui appliquent des propriétés qu'elle n'a pas. Selon une étude menée auprès d'élèves de lycée inscrits à un cours de statistiques (Mevarech, 1983), on constate que ceux-ci attribuent à la moyenne des règles de calcul qui sont celles des opérations dans des groupes : la fermeture, l'associativité, l'élément neutre et l'élément inverse (*id.*). Par exemple, 65% des élèves font la moyenne de moyennes sans tenir compte du fait que les échantillons n'étaient pas de même taille (fermeture). Ou encore, dans 80% des cas, pour faire la moyenne de trois nombres, on fait la moyenne des 2 premiers et ensuite, on prend cette dernière et on ajoute le dernier nombre avant de faire la moyenne :

$$\frac{\frac{30+40}{2}+50}{2} \neq \frac{30+40+50}{3}.$$

c'est-à-dire, 42,5 ≠ 40 ! Pour ce qui est de l'"élément neutre", 30% des sujets pensaient qu'ajouter 0 ne changeait rien à la moyenne.

Plus qu'une application de l'algorithme

Tout se complique évidemment si la moyenne est pondérée. Lorsqu'on a demandé à des étudiants américains quelle serait la moyenne finale si un étudiant avait obtenu une cote moyenne de 3,2 pour un semestre et de 3,8 pour les sept suivants, plusieurs ont tout simplement fait la moyenne des deux nombres en se justifiant ainsi : "Ce n'est pas juste mais c'est ainsi..., il faudrait avoir le détail des notes pour faire autrement !" (Pollatsek, 1981).

Lorsqu'une des données est inconnue

(4) Dans des classes de statistiques au cégep (17-19 ans), nous avons observé que certains élèves trouvaient la moyenne des effectifs qui était plus grande que la moyenne de la distribution et ils ne se rendaient pas compte de leur erreur.

(5) Voir note 3.

alors que la moyenne l'est, la situation se trouble. Le problème suivant a été présenté à des élèves du secondaire, du collégial et à des étudiants entrant à l'université.

> "Dans la classe de madame Girard, les élèves qui ont une moyenne de 6 sur 10 pour les quatre devoirs de la semaine sont dispensés de devoir pour le week-end. Après 3 devoirs, Marie a une moyenne de 5. Que doit-elle avoir comme note pour son 4^e devoir pour réussir à être dispensée de devoir pour le week-end ?" (6) (Gattuso, Mary, 1997)

Comme il n'est pas question ici d'appliquer directement un algorithme, la situation demande une meilleure compréhension. Or, les élèves de secondaire 3 ayant vu la moyenne en cours se distinguent avec 63% de réussite et surpassent les résultats des universitaires (56%), des collégiens (42%) et ceux des élèves plus jeunes n'ayant pas encore vu la leçon sur la moyenne (40%). On croirait que l'on oublie avec le temps et que la leçon fait un effet immédiat mais peu durable, si ce n'était que les collégiens qui n'ont pas mieux réussi suivaient également un cours de statistique (7). Remarquons cependant, que seulement 12,5% de tous les sujets ont utilisé la fausse propriété de la moyenne qui dit que la moyenne des moyennes est la moyenne sans tenir compte de la pondération des notes de devoirs et répondu "7" $\left(\frac{5+\langle 7\rangle}{2}=6\right)$.

En résumé, ces résultats corroborent le fait que, bien que les élèves aient une bonne idée de ce qu'est la moyenne dans des contextes simples et dans des situations simples (avec peu de données, des nombres entiers ou une moyenne non pondérée), ils ont néanmoins plus de difficulté si la moyenne est pondérée et si l'on ne peut appliquer directement un algorithme de calcul.

Quel est le problème ?

Qu'est-ce qu'il y a dans ce concept pour qu'il soit si difficile alors que, au premier abord, il semble que tout le monde sache ce qu'est une moyenne et qu'à peu près tous puissent calculer la moyenne à partir de quelques données. Or, cela ne suffit pas. C'est à travers des situations variées qu'un concept acquiert du sens pour l'apprenant. Malheureusement, bon nombre de manuels touchant ce concept se bornent à faire calculer des moyennes en ne faisant varier que les contextes. Ces exercices ne font appel qu'à l'application aveugle d'un algorithme et il ne faut donc pas s'étonner de voir des élèves trouver une moyenne de valeur supérieure aux données de la distribution. L'exploration des diverses facettes de ce concept montre que même dans ses formes élémentaires, la moyenne arithmétique est un concept mathématique très riche que l'on ne peut saisir entièrement en appliquant une formule.

Divers sens de la moyenne

Avant de proposer aux élèves une ou plusieurs variantes de la formule, que peut-on faire ? D'abord, il ne faut pas

(6) Il est courant de noter les devoirs sur 10 dans les classes du Québec.

(7) Nous ne pouvons qu'émettre une hypothèse au sujet des résultats à ce problème. Ayant examiné, plus en détails les résultats des groupes de secondaire 3, nous avons constaté qu'un de ces groupes a mieux réussi ce problème (75% versus 48,5%). L'enseignante avait utilisé des situations variées telles que présentées ici (Gattuso, Mary, 1994).

rejeter le sens habituel de la moyenne, le sens le plus courant qui "colle" à l'algorithme de calcul décrit par la formule habituelle, celui de "partage équitable", c'est-à-dire une somme également redistribuée : $\frac{\sum_{1}^{n} x_i}{n} = m$. Lors d'une expérimentation en classe (Gattuso, Mary, 1994) l'on a demandé à des élèves de secondaire de trouver la moyenne de 4 paquets de cure-dents qui contenaient respectivement 4, 6, 16, 22 cure-dents, naturellement, plusieurs ont regroupé tous les cure-dents et les ont redistribués en quatre paquets égaux. La moyenne est alors perçue comme la valeur que toutes les données auraient si elles étaient égales (Larousse, 1988). Il faudra cependant amener les élèves à constater que, malgré tout, des distributions diverses réunissant des données de valeurs différentes peuvent correspondre à une même moyenne, un même nombre de données et une même somme de données (8).

D'autres élèves ont procédé en "égalisant " les paquets : ils enlevaient des cure-dents où il y en avait plus et en rajoutaient où il y en avaient moins jusqu'à ce que les 4 paquets comprennent une même quantité de cure-dents. Pour trouver la moyenne, on peut aussi procéder par "égalisation", on peut *égaliser* en redistribuant ce qu'il y a en trop de façon à compenser pour ce qu'il y a en moins, pas à pas : "Il y en a deux de plus ici, je les mets là où il en manque...". On utilise alors une propriété importante de la moyenne, celle qui dit que la moyenne est la valeur pour laquelle la somme des écarts à cette valeur est nulle : $\sum_{1}^{n} (x_i - m) = 0$. En réalité, il y a un pas important entre l'"égalisation" et cette propriété de la moyenne car il faut de plus réaliser que le total du surplus doit être compensé par le total du déficit.

La moyenne est aussi souvent illustrée comme le point d'équilibre ou pivot d'un axe (fig. 1). Si on retrouve la propriété de la somme des écarts, on se retrouve confronté au concept physique de centre de gravité qui, un peu plus complexe, n'est pas forcément bien saisi par les élèves (9). On utilise cette représentation surtout lorsque les données sont regroupées.

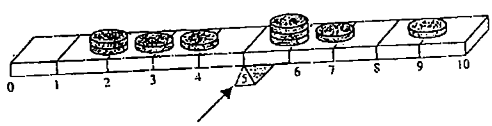


Figure 1 : Trudel, R., Antonieus, R., 1991.

La moyenne approchée de diverses façons

Regardons maintenant quelques situations en tenant compte des éléments de la réflexion précédente mais aussi en s'appuyant sur le fait que la moyenne prendra du sens pour l'apprenant s'il peut faire face à un éventail de situations très diverses qui lui permettront d'étendre son concept.

(8) Nous avons pu observer que l'implication "...**si** les données étaient égales **alors** la moyenne est la valeur que toutes les données auraient..." devient pour certains "...toutes les données sont égales à la moyenne..."

(9) Une des difficultés de ce type d'illustration est l'"immatérialité" de l'axe. Dans sa représentation mentale, l'élève doit imaginer que l'axe n'a pas de poids.

Prenons par exemple ce problème :

"La moyenne de cinq nombres différents est 4. Lorsqu'on enlève le plus grand des nombres, la moyenne des nombres qui restent est 2. Quelle est la valeur du nombre enlevé ?" (Concours Gauss 7° 1991, #16 p. 9, Concours canadien de mathématiques) (10)

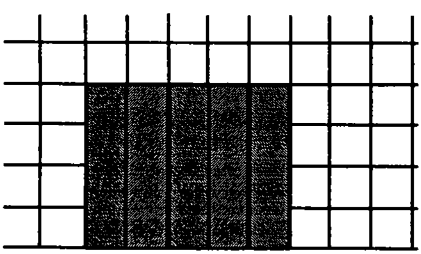
La difficulté réside dans le fait que les données de la distribution sont toutes inconnues. On peut toujours procéder algébriquement en se servant de l'algorithme somme/division :

Soit x_1. x_2, x_3. x_4, x_5 les cinq nombres, x_5 étant le plus grand.

$$\frac{X_1 + X_2 + X_3 + X_4 + X_5}{5} = 4 \rightarrow X_1 + X_2 + X_3 + X_4 + X_5 = 20 \quad (1)$$

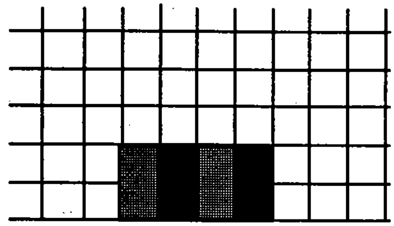
$$\frac{X_1 + X_2 + X_3 + X_4}{4} = 2 \rightarrow X_1 + X_2 + X_3 + X_4 = 8 \quad (2)$$

Ou encore s'aider du fait que la moyenne est la valeur qu'aurait chacune des données si elles étaient égales et procéder comme suit en utilisant un schéma



5 nombres dont la moyenne est 4

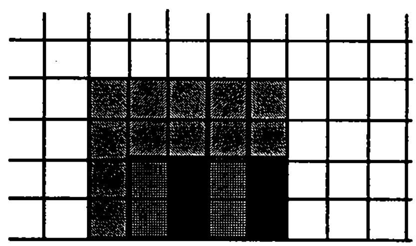
Figure 2



la moyenne des nombres qui resteent est 2

Figure 3

Si on regarde la différence entre l'aire des deux "rectangles", on peut constater que la valeur du nombre enlevé est 12.



On a enlevé 12

Figure 4

Finalement, on peut se baser sur le fait que peu importe la distribution, pour une moyenne donnée, la somme des données restera la même. L'hypothèse de l'égalité de toutes les données n'est pas absolument nécessaire ici. "Si la moyenne des 5 nombres est 4, alors la somme des nombres est 4×5 puisque la moyenne 4 est obtenue en partageant le total en 5. De même si la moyenne de 4 nombres est 2 alors la somme des nombres restants est $2 \times 4 = 8$. Il faut aussi voir que la somme des cinq nombres étant 20 et la somme de 4 de ces nombres étant 8, le cinquième est nécessairement 12." Ce raisonnement est essentiellement ce qui est fait algébriquement plus haut mais sans symbolisme.

Cette façon de travailler peut se trans-

(10) Le concours Gauss 7° s'adresse aux élèves de 12 ans -7e année.

férer à des situations plus complexes. Cette stratégie qui fait appel à la différence des sommes a été une des plus utilisées (23,8%) et des plus réussies par les élèves en réponse à la question "Les devoirs" citée plus haut (Gattuso *et al.*, 1997). En effet, plusieurs d'entre eux ont dit : "Il faut en tout 24 points pour les 4 devoirs pour avoir 6 de moyenne, si Marie a une moyenne de 5 pour les trois premiers devoirs, elle a déjà 15 points. Donc, 24 – 15 = 9, c'est ce qu'elle doit obtenir pour le dernier devoir". On voit qu'une même situation peut être travaillée en faisant appel à divers "sens" de la moyenne. La solution algébrique se base sur l'algorithme, la solution graphique sur celui de l'*égalisation* des données et la dernière sur la somme des données.

Par ailleurs, comme nous l'avons mentionné plus haut, dans l'*égalisation*, se cache la propriété qui dit que la moyenne est la valeur pour laquelle la somme des écarts à cette valeur est nulle : $\sum_{1}^{n}(x_i - m) = 0$. Cette propriété peut parfois être fort utile. Regardons l'exemple suivant :

> Massamba reçoit un ensemble de quatre notes. Si la moyenne des deux premières notes est 50, la moyenne des deuxième et troisième notes est 75 et la moyenne des troisième et quatrième notes est 70, alors quelle est la moyenne des première et quatrième notes ? (Concours Cayley, 10° 1991, #20 p33, Concours canadien de mathématiques) [11].

(11) Le concous Cayley 10° s'adresse aux élèves de 15 ans.

Encore ici, on peut procéder algébriquement :

$$\begin{array}{lll}
x_1 + x_2 = & 100 & (1) \\
x_2 + x_3 = & 150 & (2) \\
x_3 + x_4 = & 140 & (3) \\
x_1 - x_3 = & -50 & (1) - (2) = (4) \\
x_1 + x_4 = & 90 & (4) + (3) = (5) \\
\dfrac{x_1 + x_4}{2} = & 45 &
\end{array}$$

Ceci suppose tout de même trois équations et quatre inconnues. La plupart des élèves ne réussissent pas à mener à terme cette résolution. Toutefois, un raisonnement utilisant la propriété de la somme des écarts, soutenu par un schéma, conduira au même résultat.

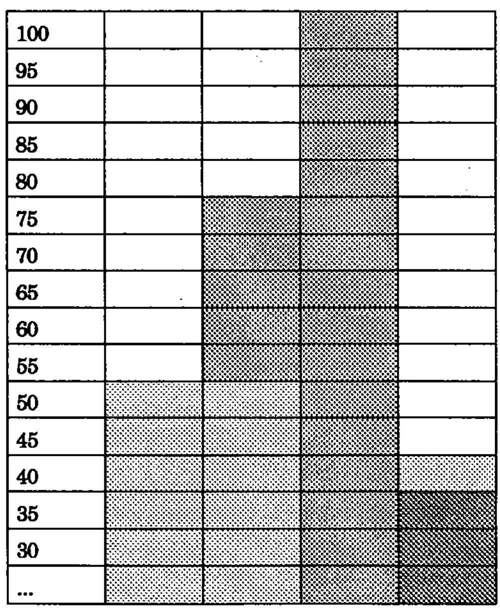


Figure 5

En supposant que les deux premières notes soient 50, on voit sur le schéma que, pour que la moyenne de la deuxième et troisième note soit 75, il faudra, pour compenser l'écart de -25 qu'il y a entre la deuxième note, 50 et la moyenne, 75, que la troisième note soit 75 + 25 = 100. Comme la moyenne des troisième et quatrième notes est 70, pour compenser l'écart de +30 qu'il y a entre la troisième note et la moyenne, il faudra un écart de –30, donc la quatrième note sera 70 – 30 = 40. Et finalement, la moyenne de la première et quatrième note sera (50 + 40)/2 = 45. Il y a cependant lieu de se demander ce qui se passerait si l'hypothèse de départ était différente. C'est un très bon exercice pour les élèves ; en faisant divers essais, ils se rendront compte qu'ils arrivent toujours au même résultat. En déplaçant des languettes sur un transparent représentant à peu près le schéma ci-haut, on peut les faire réfléchir sur ce jeu d'écart et montrer (en déplaçant les languettes) que dès que l'on change une valeur, le reste doit s'ajuster si on garde les mêmes moyennes. Tout est lié. La moyenne est liée à chacune des valeurs de la distribution, c'est une autre de ses propriétés. Jouer sur une donnée a un impact sur les autres, la moyenne étant maintenue. Jouer sur une donnée a un impact sur la moyenne, les autres données étant fixées. Une autre considération intéressante est de regarder les valeurs limites que peut prendre chacune des notes. Peuvent-elles toutes varier de 0 à 100 ?

Signalons que ce problème ne comporte aucune donnée brute, on ne connaît pas les notes elles-mêmes, mais seulement des moyennes partielles. Alors, est-il possible de "calculer" la moyenne des quatre notes en utilisant l'algorithme "somme-division" ?

On peut le faire algébriquement, les notes étant les inconnues :

$$\frac{x_1+x_2}{2}=50 \rightarrow x_1=100-x_2$$

$$\frac{x_2+x_3}{2}=75 \rightarrow x_2=150-x_3$$

$$\frac{x_3+x_4}{2}=70 \rightarrow x_3=140-x_4$$

$$\rightarrow x_4=140-x_1$$

de là, on trouve

$x_1 + x_2 + x_3 + x_4 = 480 - (x_1 + x_2 + x_3 + x_4)$

et $\quad x_1 + x_2 + x_3 + x_4 = 240$

la moyenne est donc 60 puisque la somme des quatre notes est 240.

Ou encore, en supposant comme plus haut que les deux premières notes soient 50, on trouve que la troisième est 100 et la quatrième 40. Ensuite, on peut trouver la moyenne des quatre notes en jouant sur les écarts. En posant comme moyenne hypothétique 50, on constate que l'écart entre la première et deuxième note et la moyenne est 0. L'écart entre la troisième note, 100 et la moyenne, 50 est de +50 et pour la quatrième note il est de –10. La somme des écarts n'est pas nulle (40) donc 50 n'est pas la moyenne. Il faut prendre une moyenne plus grande. Reprenant la même procédure avec 60, on trouve 0 pour la somme des écarts à la moyenne. Donc la moyenne est 60.

Moyenne hypothétique	50	60
Note 1 : 50	0	–10
Note 1 : 50	0	–10
Note 3 : 100	+50	+40
Note 4 : 40	–10	–20
Somme des écarts	**+40**	**0**

Évidemment on aurait pu, une fois les notes admises, les additionner et les diviser par 4... ! Cela est simple à faire, une fois les notes trouvées évidemment ! Mais est-ce toujours possible ? Voici un exemple tiré de *Carrousel mathématique, secondaire 3*, vol. 2, p. 266 :

A un groupe d'élèves ayant une moyenne de 72%, on ajoute deux élèves ayant des notes de 76% et 68%.
Peut-on calculer la nouvelle moyenne ? Si oui, calcule-la. Sinon, indique pourquoi.

Ici, on ne connaît ni les données, ni le nombre de données. On serait tenté de répondre : "non !" mais on peut essayer une résolution algébrique.

$$\frac{72 \cdot x + 144}{x+2} = m$$

$$\frac{72 \cdot x + 144}{x+2} = m$$

$$\frac{72(x+2)}{x+2} = m \to m = 72$$

On constate que la moyenne n'a pas changé. Est-ce toujours vrai ? Répondons à partir d'une illustration.

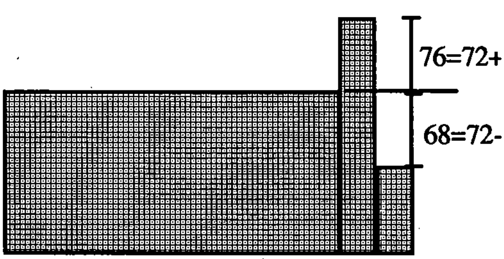


Figure 6

En utilisant le fait que l'on peut voir la moyenne comme la valeur qu'auraient toutes les données si elles étaient égales et que, pour les valeurs ajoutées, la somme des écarts à 72 est de zéro, on peut conclure que la nouvelle moyenne [12] est 72. Il faut insister sur le fait que ce cas est particulier car la moyenne des valeurs ajoutées est 72. Tout ajout de valeurs dont la moyenne est égale à la moyenne initiale de modifie pas cette dernière.

Estimation de la moyenne de données pondérées

Travailler avec les écarts permet d'estimer une moyenne à partir d'une représentation graphique et ceci même dans les cas de moyenne pondérée. Dans ces cas, cependant, il est souvent plus utile de faire appel à une représentation statistique comme le diagramme à bâton ou l'histogramme. Néanmoins, il faut être attentif car ici nous ne faisons plus appel tout à fait au même sens de la moyenne. Même si nous utilisons le fait que la somme des écarts à la moyenne doive être nulle, cette façon de travailler mise aussi sur une compréhension au moins implicite du point d'équilibre ou du pivot représentée plus haut (Figure 1) plutôt que sur le fait que la moyenne est la valeur qu'auraient toutes les données si elles étaient égales. Illustrons nos propos par un exemple (St-Hilaire, 1997).

variable X	fréquences
10	3
20	1
30	1
40	2

(12) Et que la moyenne ne changera pas à la condition que, pour les notes ajoutées, la somme des écarts à la moyenne, 72, soit nulle.

Si nous cherchons à estimer la moyenne de ces données sans faire de calculs en travaillant avec une représentation graphique,

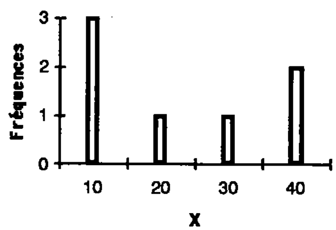


Figure 7

on peut remplacer sur le diagramme en bâton (Figure 7), les deux données valant 40 par deux données valant 30. On obtient le deuxième graphique (Figure 8).

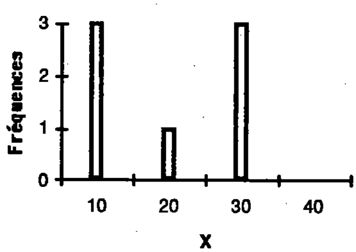


Figure 8

La moyenne est alors plus petite que la moyenne cherchée à l'origine et elle est de 20 puisque à cause de la symétrie du graphique le point d'équilibre est le point milieu. Par conséquent, la moyenne originale est plus grande que 20 car l'on a modifié les données en diminuant la valeur de deux d'entre elles.

Si, au lieu de remplacer deux données valant 40 par deux données valant 30, on enlève une donnée valant 10, on obtient le diagramme qui suit.

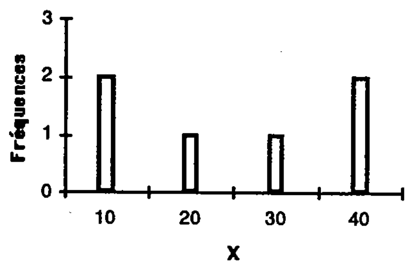


Figure 9

Ce faisant, le point d'équilibre se déplace vers la droite. La moyenne, sur ce diagramme est 25, le point milieu, car le graphique est symétrique par rapport à un axe vertical passant par 25. On peut également faire remarquer que si on enlève un nombre inférieur à la moyenne, on augmente la moyenne. Par conséquent, la moyenne originale est plus petite que 25. On conclut que la moyenne cherchée se trouve entre 20 et 25. En faisant le calcul, on trouve 22,86.

Ce jeu de graphiques n'est pas aussi simple qu'il peut paraître. Il faut d'abord comprendre que le point d'équilibre se trouve au milieu si la distribution est symétrique. Le graphique le facilite. Ensuite, on doit prendre conscience que toute modification d'une donnée change la distribution et par conséquent la moyenne. Encore ici, le graphique l'illustre bien. Finalement, on peut anticiper les effets des modifications d'une donnée sur la moyenne. Si l'on diminue les valeurs, la moyenne diminue et vice versa. Il n'est pas aussi évident qu'enlever une petite valeur fera augmenter la moyenne, mais, encore là, l'utilisation de représentations graphiques devrait aider l'élève à s'en convaincre.

Cependant, il faut être prudent. Introduire au même moment les représentations où les données sont regroupées et les autres où les données sont brutes risque d'embrouiller les élèves. Dans ce dernier cas, on tente d'obtenir un rectangle et de rendre les données *égales* (13), avec les données pondérées, on tente d'avoir une représentation symétrique dans le but d'estimer la moyenne, point d'équilibre, qui, dans ce cas, se situe au milieu du graphique sur l'axe horizontal. Encore ici, il peut y avoir confusion car on favorise la conception parfois inadéquate que la moyenne est au milieu. De plus, on fait une analogie avec le pivot ou la balance alors que les principes en jeu ne sont pas nécessairement déjà acquis par les élèves. Il sera peut-être nécessaire d'introduire ces principes en commençant par des activités à base de manipulations afin de développer chez les élèves la compréhension de l'équilibre. Rajoutons que ce "jeu de graphiques" n'est pas nécessairement très efficace, mais il fait réfléchir et s'appuie sur le raisonnement et les propriétés de la moyenne qui valent la peine d'être activés.

En guise de conclusion

Nous pouvons constater que la moyenne n'est pas simple à apprendre et elle n'est pas simple à enseigner. Avant d'amorcer le travail avec les élèves, il est important, au point de départ, de sonder leurs représentations pour les confronter en présentant une situation semblable à celle-ci :

> – J'ai lu dans le journal que : *"Les adolescents français regardent la télévision en moyenne 22 heures par semaine."* Qu'est-ce que ça veut dire : *en moyenne 22 heures par semaine* ? Est-ce que ça veut dire que tous les enfants regardent 22 heures de télévision par semaine ?

En réponse à cette question, les élèves révèlent leurs représentations au sujet de la moyenne : "À peu près"... "Habituellement"... "Le plus souvent"... On écrit leurs réponses au tableau pour y revenir afin de les comparer avec la définition de la moyenne qu'ils auront formulé par la suite.

Il faut poursuivre et confronter les idées préconçues des élèves. Pour cela, au lieu de demander de trouver une moyenne, à partir d'une moyenne donnée, on peut demander de trouver les données tout en posant des contraintes bien choisies.

> – Caroline, Vincent, Juliette et Yannick ont en poche en moyenne 100 francs chacun. Est-ce possible qu'aucun d'entre eux n'ait 100 francs exactement ? Distribuez vos montants et essayez de trouver une possibilité où personne n'a exactement 100 francs.

Cette contrainte (personne n'a 100 francs) vise à souligner que la moyenne n'est pas nécessairement une des données. De plus, les élèves peuvent trouver diverses solutions et ils constateront que diverses distributions ont la même moyenne. Il est nécessaire d'insister sur ces deux derniers points. Ensuite on peut demander s'il est possible qu'un seul enfant ait 100 F ? Et deux enfants... ? Et exactement trois... ? Et finalement, est-ce que tous peuvent avoir 100 F ? Lorsqu'un seul enfant a 100 F, la moyenne n'est pas la valeur la plus fréquente et n'est pas non plus celle du milieu. Si deux ont 100 F, on devra trouver des valeurs équidistantes de 100 F, renforçant ainsi l'idée que la somme des écarts à la moyenne doit être nulle. Pour cette raison également, il est impossible qu'exactement 3 enfants aient 100 F. Et si quatre enfants ont 100 F la moyenne est la valeur la plus fréquente dans ce cas !

Les variations possibles des énoncés

(13) Voir la figure 6.

sont nombreuses. On peut demander de donner un exemple de distribution ou la moyenne est (ou n'est pas) au milieu des données (il y en a autant avant et après) ou au milieu des valeurs (au milieu entre la plus petite et la plus grande). Par la suite, on pourra travailler avec de plus grands nombres et passer à des situations plus complexes. On peut chercher une ou plusieurs données qui sont simples ou pondérées. L'effectif total peut être inconnu ou encore, il est possible que l'on ne connaisse que la variation de la moyenne. Il est également nécessaire de se questionner sur ce qui se passe pour la moyenne si on augmente également toutes les données. Par exemple, si la moyenne des notes de l'examen est de 67, que se passe-t-il si le professeur donne un bonus de 7 points à tout le monde ? Doit-il recalculer sa moyenne ? Peut-on trouver une autre façon de procéder ?

Nous avons tenté de vous convaincre que la moyenne est un concept beaucoup plus riche qu'il n'y paraît au premier regard. L'exploration que nous avons présentée n'est qu'une amorce. La moyenne intervient dans plusieurs autres contextes. On pense entre autres à la vitesse moyenne et au centre de masse en physique. N'oublions pas les problèmes que l'on qualifie souvent de "problèmes de mélanges" qui eux aussi sont des problèmes qu'il serait plus facile d'aborder sous l'angle de la moyenne. Évidemment, la moyenne est un concept essentiel à l'inférence statistique et aux statistiques en général.

Pour terminer, disons que nous avons voulu montrer, en prenant comme exemple la moyenne, que les statistiques sont plus qu'une branche dérivée des mathématiques ou encore une de ses applications. S'il est vrai que les connaissances mathématiques sont indispensables pour l'apprentissage des statistiques, l'étude des ces dernières devrait être intégrée plus étroitement à l'étude des mathématiques car leur apport est très enrichissant.

RÉFÉRENCES

BOLON, J. (1992) : "L'enseignement des décimaux à l'école élémentaire", *Grand N. 52,* 49-79.

BROUSSEAU, G. (1981) : "Problèmes de didactique des décimaux.". *Recherches en Didactique des Mathématiques (2),* 1, 37-127.

BROUSSEAU, G. (1988) : "Représentations et didactique du sens de la division" *in* Vergnaud, Gérard, Brousseau, Guy, Hulin, Michel. (Eds.) *Didactique et acquisition des connaissances scientifiques. Actes du Colloque de Sèvres qui s'est tenu au CIEP en mai 1987,* Grenoble : Editions La Pensée Sauvage, 47-64.

GATTUSO L., Mary C. (1994) : *La moyenne.* Leçon réalisée dans une classe de secondaire III, d'après une préparation de cours d'un étudiant de Didactique I (MAT2023), Document vidéoscopique.

GATTUSO, L. (1997) : "La moyenne, un concept élémentaire ?", *Ortho...graphe plus.*

GATTUSO, L., MARY, C. (1996) : "Development of concepts of the arithmetic average from high school to university", *in* Luis Puig & Angel Gutiérrez (eds.), *Proceeding of the 20th International Conference for the Psychology of Mathematics Education*, Valence, Espagne, 401-408.

GATTUSO, L., Mary, C. (1997) : "La moyenne, un concept évident ?", *Bulletin de l'AMQ*. octobre, 10-19.

LEON, M., ZAWOJEWSKI, J. (1990) : "Use of Arithmetic Mean : An Investigation of Four Properties Issues and Preliminary Results", *in* David Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics*, Dunedin, New Zealand, 302-306.

MAURIN, R. (1981) : "L'enseignement des fractions à l'école élémentaire", *Bulletin de l'APMEP. 327*, pp. 41-47

MEVARECH, Z.R. (1983) : "A deep structure model of students' statistical misconceptions", *Educational Studies in Mathematics, 14,* 415-429.

Mokros, J., Russell, S. (1991) : "Toward an understanding of mean as 'balance point'", *in* R.G. Undehill (Ed.) : *Proceedings of the Thirteenth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, PME-XI*. Blacksburg : Virginia Polytechnic Institute and State University, 189-195.

MOKROS, J., RUSSELL, S. (1995) : "Children's concepts of average and representativeness", *Journal for Research in Manthematics Education, 26,* (1) 20 39.

PERRIN, M.-J. (1986) : "Représentation des fractions et des nombres décimaux chez des élèves de CM2 et du collège", *Petit X,* 10, 5-29

POLLATSEK, A., LIMA, S., WELL, A.D. (1981) : "Concept or computation : Students' understanding of the mean", *Educational Studies in Mathematics, 12,* 191-204.

RUSSELL, S., MOKROS, J. (1990) : "What's Typical ? Children's and Teachers' Ideas About Average", *in* David Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics*, Dunedin, New Zealand, 307-313.

ST-HILAIRE, C. (1997)
http ://www.virtuel.collegebdeb.qc.camaths/claudesh/stats/prestat.htm

STRAUSS, S., BICHLER, E. (1988) : "The development of children's concepts of the arithmetic average", *Journal for Research in Manthematics Education, 19,* 64-80.

TRUDEL,r., ANTONIEUS, R. (1991) : *Méthodes quantitatives appliquées aux sciences humaines*, Montréal : Les Éditions de la Chenelière.

VERGNAUD, G. (1990) : "Développement et fonctionnement cognitifs dans le champ conceptuel des structures additives", *in* Netchine-Grynberg, Gaby, (Ed) *Développement et fonctionnement cognitifs chez l'enfant*, Paris : Presses Universitaires de France, 261-277.

VERGNAUD, G. RICCO, G. ROUCHIER, A. *et al* (1978) : "Quelles connaissances les enfants de sixième ont-ils des 'structures multiplicatives' élémentaires ? Un sondage", *Bulletin de l'APMEP, 313,* 331-357.