St2 - LE MEILLEUR DES NOMBRES

TI-82 STATS – TI-83 Plus – TI-84 Plus

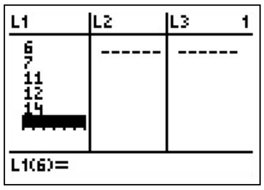
Mots-clés : médiane, moyenne, écart-type, distance, minimum, valeur absolue.

1. Objectifs

À partir du problème de minimisation « Quel est le nombre le plus proche des valeurs d'une série statistique ? », retrouver la médiane (solution lorsque l'on prend la somme des valeurs absolues des écarts) et la moyenne arithmétique (solution lorsque l'on prend la somme des carrés des écarts).
Légitimer la définition d'écart type.

2. Mise en œuvre et commentaires

Sur une droite graduée, on place cinq nombres, par exemple 6, 7, 11, 12 et 14.



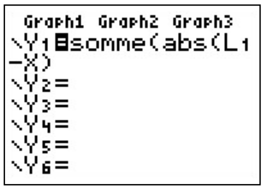
écran 1

Problème

Comment peut-on représenter (ou résumer) ces cinq nombres par un seul, ou encore quel nombre choisir pour qu'il soit le plus « **proche** » possible de ces cinq nombres ?

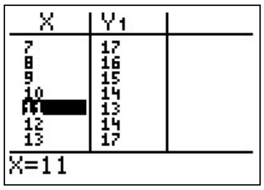
Partie A

Comme définition de distance d'un nombre à une série de valeurs, on prend la somme des valeurs absolues des écarts (distance entre 2 nombres rencontrée en seconde).



écran 2

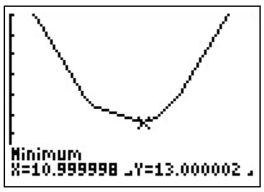
Les nombres sont rangés dans la liste L1 (*écran* 1). On note $D(x)$ cette distance. $D(x)$ peut donc être saisie dans l'éditeur de courbes par la formule Y1=somme(abs(L1 – X)) (*écran* 2).



écran 3

1) Recherche

On conjecture le minimum de cette fonction (et la valeur de x pour laquelle il est atteint) à partir d'une table de valeurs (*écran* 3), puis à l'aide de sa représentation graphique (*écran* 4). On remarque en particulier que ce n'est pas la moyenne qui minimise la somme D.



écran 4

Remarques :

- *D est une fonction affine par morceaux. Cet aspect, plus facilement visible sur un ordinateur, peut être expliqué avec une série de deux valeurs seulement (par disjonction des cas), mais ce n'est pas un objectif de l'activité.*
- *Les valeurs renvoyées par la calculatrice, comme toujours, sont a priori des approximations.*

En ajoutant une valeur dans la série (ce qui fait qu'elle en contient un **nombre pair**), le minimum de la fonction D est atteint non plus pour une valeur mais pour un **intervalle de valeurs** (D étant constante sur cet intervalle) : la solution au problème n'est plus unique.

2) Démonstration

■ **1^er^ cas :** le nombre de valeurs est impair (5 dans notre exemple)

On montre que tout nombre de [6 ; 14] minimise la somme $|6 - x| + |14 - x|$ (6 et 14 sont les éléments extrêmes de la liste). Le dessin sur une droite graduée en donne une vision simple :

$$\text{pour tout } x \text{ de } [6\ ;\ 14],\ |6 - x| + |14 - x| = d(14\ ;\ 6) = 8.$$

De même tout nombre de [7 ; 12] minimise la somme $|7-x|+|12-x|$. Comme [7 ; 12] $\subset$ [6 ; 14], on en déduit que tout nombre de [7 ; 12] minimise aussi la somme $|6-x|+|14-x|+|7-x|+|12-x|$.
Enfin, comme seul 11 minimise $|11 - x|$, on en déduit que 11 minimise $D(x)$. Mais plus que cela, on établit que, dans le cas d'un nombre impair de valeurs, on arrive par ce raisonnement à montrer que la solution est le nombre qui sépare l'effectif en deux parts égales : c'est donc la **médiane**.

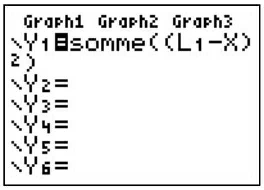
■ **2nd cas :** le nombre de valeurs est pair (6 dans notre exemple)
Le raisonnement précédent montre que l'on arrive à un intervalle final (le plus petit emboîté dans les autres ; ici, l'intervalle [8 ; 11]).
Dans le cas d'un nombre pair de valeurs, on convient que la médiane est le milieu de l'intervalle solution.

Partie B

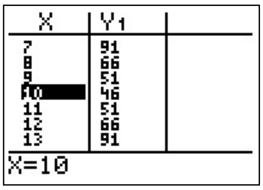
Comme définition de distance d'un nombre à une série de valeurs, on prend cette fois la somme des carrés des écarts. On note $P(x)$ cette distance. $P(x)$ peut donc être saisie dans l'éditeur de courbes par la formule Y1=somme((L1–X)^2).

1) Recherche

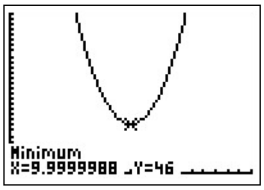
On suit le même plan de travail (*écrans* 5 à 7).



écran 5



écran 6



écran 7

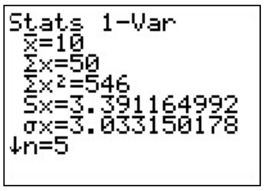
Cette fois, il semble bien que la **moyenne** minimise P.

2) Démonstration

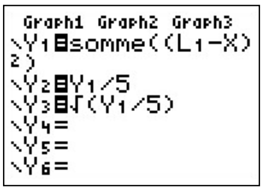
$P(x)$ est une fonction trinôme. Après développement, on peut déterminer son minimum en appliquant le cours sur le second degré si les élèves en disposent. Dans le cas contraire (en seconde ou 1ère L), la factorisation de $P(x)-P(\bar{x})$ permet de conclure.

Définitions de variance et écart type.

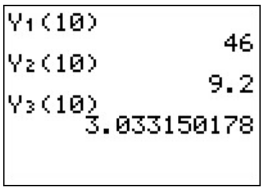
Après avoir remarqué que $P(x) \neq (D(x))^2$, on définit la **variance** de la série par $v(\bar{x}) = \dfrac{P(\bar{x})}{5}$ où 5 est l'effectif de la série et $\bar{x}$ la valeur trouvée qui minimise P, puis l'**écart type** par $s = \sqrt{v(\bar{x})}$. On peut vérifier que la commande intégrée de la calculatrice donne bien le même résultat pour ces paramètres (dans **STAT** sélectionner le menu **CALC**, puis **1: Stats 1-Var**) et en superposant les fonctions P, v et s, on peut vérifier graphiquement qu'elle ont un minimum pour la même valeur $\bar{x}$ qui est la moyenne arithmétique de la série.



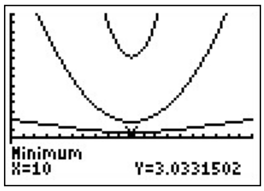
écran 8



écran 9



écran 10



écran 11

Pour écrire **Y1** et **Y2** (*écrans* 9 et 11): dans **VARS**, ouvrir le menu **Y-VARS**, puis **1: Fonction** et sélectionner **Y1** ou **Y2**.

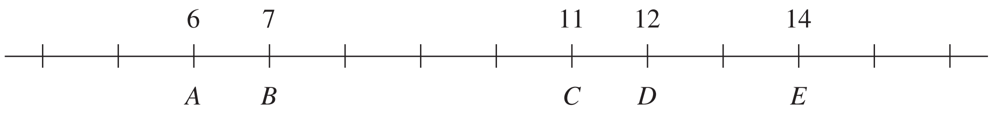
Les démonstrations ont été faites dans le cadre d'un exemple numérique très simple. Cela permet aux élèves de bien identifier les éléments importants sur lesquels elles reposent, et, dans le cas où l'on voudrait faire les démonstrations générales, cela permet de mieux les comprendre.

Nom .... Classe ....

St2 - LE MEILLEUR DES NOMBRES

TI-82 STATS – TI-83 Plus – TI-84 Plus

Sur une droite graduée, on a représenté les cinq nombres 6, 7, 11, 12 et 14 :



Problème

Comment peut-on résumer ces cinq nombres par un seul ? On souhaite choisir celui qui est le plus proche des 5 nombres. Comment le déterminer ?

Partie A : Le plus proche en considérant la distance : $d(x\ ;a) = |x - a|$

Soit M un point de la droite, d'abscisse x.

On note $D(x)$ la somme des distances MA, MB, MC, MD et ME.

Exprimer $D(x)$ en fonction de x.

On cherche un nombre x pour lequel $D(x)$ est minimale.

1) Recherche numérique et graphique à l'aide de la calculatrice

Dans l'éditeur de listes, entrer les cinq nombres dans L1 (**STAT 1: Edite…**).

Dans l'éditeur de courbes (**Y=**), en Y1, saisir l'expression : somme(abs(L1 – X)).

La commande somme se trouve dans **2nd [LIST]**, menu **MATH**, et **abs** dans **MATH** menu **NUM**.

Dans Y1 on a ainsi défini la fonction D.

a) À l'aide de tableaux de valeurs de la fonction D (**2nd [TABLE]** configuré dans **2nd [TBLSET]**), rechercher pour quelle valeur de x la somme D est minimale.

Cette valeur est-elle la moyenne de la série ?

b) À l'aide de l'éditeur de courbes, afficher la courbe représentative de D dans une fenêtre adaptée.

Avec la commande **TRACE** (ou bien avec la commande **2nd [CALC] 3: minimum**), retrouver pour quelle valeur de x la somme D est minimale.

c) Faire la même recherche en ajoutant 8 à la liste des nombres.

2) Démonstration

Enlever le chiffre 8 de la liste.

a) M étant un point d'une droite graduée, où doit se situer le point M pour que la somme des distances $MA + ME$ soit minimale ?

En déduire l'intervalle dans lequel on doit prendre le nombre x afin que la somme $s_1 = |6 - x| + |14 - x|$ soit minimale.

b) Où doit se situer le point M pour que la somme des distances $MB + MD$ soit minimale ?

En déduire l'intervalle dans lequel on doit prendre le nombre x afin que la somme $s_2 = |7 - x| + |12 - x|$ soit minimale.

En déduire l'intervalle dans lequel on doit prendre le nombre x afin que la somme :

$$s_1 + s_2 = |6 - x| + |14 - x| + |7 - x| + |12 - x| \text{ soit minimale.}$$

c) En quel nombre x la valeur $|11 - x|$ est-elle minimale ?

d) De l'étude précédente, déduire le nombre m tel que la somme $D(x)$ soit minimale.

Que représente ce nombre pour la série statistique $\{6\ ;7\ ;11\ ;12\ ;14\}$?

e) On note $\bar{x}$ la valeur moyenne de cette série. Vérifier par le calcul que $D(m) < D(\bar{x})$.

f) Refaire la démonstration en ajoutant 8 à la liste.

Partie B : Le plus proche en considérant le carré de la distance : $[d(x\ ;a)]^2 = (x - a)^2$

On note $P(x)$ la somme des carrés des distances MA, MB, MC, MD et ME.
Exprimer $P(x)$ en fonction de x.
On cherche cette fois le nombre x pour lequel $P(x)$ est minimale.

1) Recherches numérique et graphique à l'aide de la calculatrice

En procédant comme précédemment, chercher à l'aide de la calculatrice quel semble être ce minimum.

2) Recherche de la démonstration du résultat

a) Comparer $P(m)$ et $P(\bar{x})$, où m est la valeur trouvée à la **Partie A**.

b) Développer $P(x)$ et donner sa forme réduite, ordonnée suivant les puissances de x.

c) Étudier le signe de $P(x) - P(\bar{x})$. En déduire que P admet un minimum et préciser en quelle valeur de x.

3) Définition de la variance et de l'écart type

a) En calculant $\big(D(x)\big)^2$ retrouve-t-on $P(x)$? Expliquer pourquoi.
Avec la calculatrice, retrouver ce résultat graphiquement.
Pour tenir compte de la taille de la liste de nombres étudiée, on divise la somme minimale $P(\bar{x})$ par le nombre de termes. On obtient donc le nombre $\frac{P(\bar{x})}{5}$. Ce nombre est appelé *variance* de la série de valeurs ; la racine carrée de la variance est appelé *écart type* de la série :

• variance : $v = \frac{P(\bar{x})}{5}$; • écart type : $s = \sqrt{v} = \sqrt{\frac{P(\bar{x})}{5}}$.

b) Dans un même repère, représenter les trois fonctions :

$$x \mapsto P(x)\ ,\ x \mapsto \frac{P(\bar{x})}{5} \text{ et } x \mapsto \sqrt{\frac{P(\bar{x})}{5}}\ .$$

Pour quelle valeur atteignent-elles leur minimum ?
Que représente le minimum de la deuxièmefonction ? et celui de la troisième fonction ?