STATISTIQUES À UNE VARIABLE

Notations :

$x_1, x_2, \dots, x_p$ sont les valeurs ou les centres des classes si ces valeurs sont regroupées en classe.

$n_1, n_2, \dots, n_p$ sont les effectifs respectifs des valeurs $x_1, x_2, \dots, x_p$.

$f_1, f_2, \dots, f_p$ sont les fréquences respectives des valeurs $x_1, x_2, \dots, x_p$.

N est l'effectif total : $N = n_1 + n_2 + \dots + n_p$

I. Moyenne, variance et écart type :

1°) Moyenne :

Définition :

La moyenne d'une série statistique est le nombre notée $\overline{x}$ et a pour valeur :

$$\overline{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_p x_p}{N} = \frac{1}{N} \sum_{i=1}^{p} n_i x_i .$$

ou, à l'aide des fréquences :

$$\overline{x} = f_1 x_1 + f_2 x_2 + \dots + f_p x_p = \sum_{i=1}^{p} f_i x_i .$$

Exemples :

1. caractère quantitatif discret :

Le tableau ci-dessous donne la répartition des élèves d'un lycée suivant l'âge :

Âge	14	15	16	17	18	19	20
effectif	130	204	270	316	198	77	14

L'âge moyen est de : $\overline{x} = \frac{14\times130+15\times204+16\times270+17\times316+18\times198+19\times77+20\times14}{1209} \approx 16,44$.

2. caractère quantitatif continu :

Le tableau ci-dessous donne la répartition des tailles en cm de 60 bébés à la naissance :

Taille	[42 ; 46[	[46 ; 50[	[50 ; 54[	[54 ; 58[
effectif	13	17	23	7

Dans ce cas, il faut prendre le centre de chaque classe pour faire la moyenne :

La taille moyenne est de : $\overline{x} = \frac{44\times13+48\times17+52\times23+56\times7}{60} = 49,6$.

2°) Variance et écart type :

Définition :

La variance, notée V, d'une série statistique est définie par :

$$V = \frac{1}{N}\left[n_1(x_1-\overline{x})^2 + n_2(x_2-\overline{x})^2 + \dots + n_p(x_p-\overline{x})^2\right] = \frac{1}{N}\sum_{i=1}^{p} n_i(x_i-\overline{x})^2$$ où $\overline{x}$ est la moyenne de la série.

Définition :

L'écart type, noté σ, est la racine carrée de la variance : $\sigma = \sqrt{V}$.

Remarques :

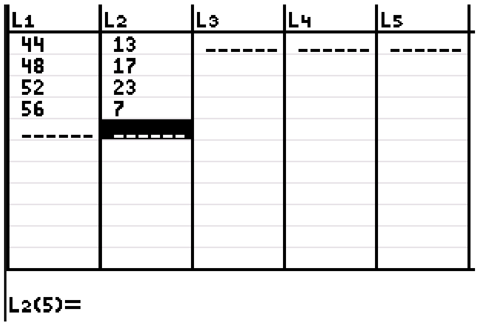
L'écart type mesure la dispersion des valeurs de la série autour de la moyenne : plus l'écart type est petit, plus les valeurs de la série sont concentrées autour de la moyenne.

L'écart type à la même unité que les valeurs de la série étudiée.

Méthode :

Pour trouver l'écart-type d'une série statistique, on utilise la calculatrice. Sur la TI, on commence par appuyer sur la touche 

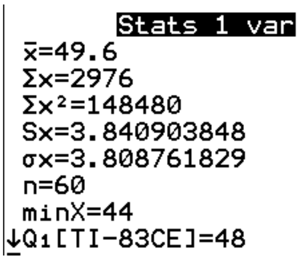
, puis dans l'onglet « edit » on sélectionne 1:Modifier… . On saisit les valeurs de la séries dans la colonne L_1 et les effectifs ou les fréquences dans L_2. On obtient ceci si on saisit les données de l'exemple sur la taille des bébés :

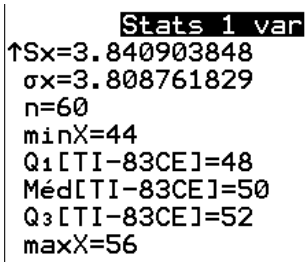


On appuie sur la touche 

pour revenir au menu statistique, puis, dans l'onglet CALC on sélectionne 1:Stats 1 Var puis on saisit :

Pour faire L1, on tape 

On obtient alors 

puis 

en appuyant sur la flèche qui descend.

$\bar{x}$ est la moyenne et σx est l'écart-type.

Si l'on veut la variance, il suffit de mettre l'écart-type au carré.

II. Médiane et quartiles :

1°) Médiane :

Définition :

C'est la valeur **Me** pour laquelle la **moitié au moins** des individus de la population a une valeur pour le caractère étudié **supérieure ou égale** à la médiane et **la moitié au moins** a une valeur **inférieure ou égale** à la médiane.

Exemples :

- Si N est impair, la médiane est sa valeur centrale :

Notes de 11 élèves : 5 ; 6 ; 7 ; 7 ; 9 ; **11** ; 12 ; 12 ; 13 ; 14 ; 15. La médiane est la $6^{\text{ème}}$ valeur, soit Me = 11

- Si N est pair, la médiane est la demi-somme (moyenne) des deux valeurs centrales :

Notes de 10 élèves : 5 ; 6 ; 7 ; 7 ; **9 ; 12** ; 12 ; 13 ; 14 ; 15. La médiane est la moyenne entre la $5^{\text{ème}}$ et la $6^{\text{ème}}$ valeur, soit Me = $\dfrac{9+12}{2}$ = 10,5.

2°) Quartiles :

Définitions :

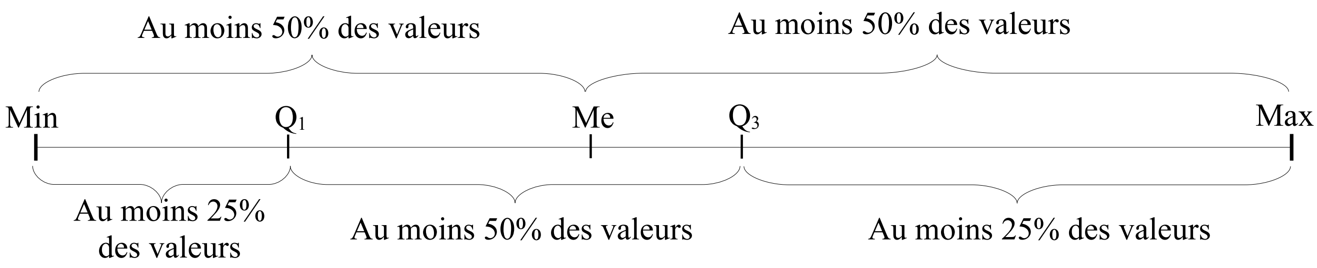
Le **premier quartile** d'une série, noté $\mathbf{Q_1}$, est la plus petite valeur de la série pour laquelle au moins **un quart (25%)** des données sont inférieures ou égales à $\mathbf{Q_1}$.

Le **troisième quartile** d'une série, noté $\mathbf{Q_3}$, est la plus petite valeur de la série pour laquelle au moins **trois quarts (75%)** des données sont inférieures ou égales à $\mathbf{Q_3}$.

Définitions :

On appelle intervalle interquartile, l'intervalle $[\mathbf{Q_1} \,;\, \mathbf{Q_3}]$.

On appelle l'écart interquartile, noté **Q**, l'amplitude de l'intervalle $[\mathbf{Q_1} \,;\, \mathbf{Q_3}]$ c'est à dire le nombre $\mathbf{Q} = \mathbf{Q_3} - \mathbf{Q_1}$.



Exemple :

Caractère quantitatif discret (retour aux notes des élèves)

- Notes de 11 élèves : 5 ; 6 ; 7 ; 7 ; 9 ; 11 ; 12 ; 12 ; 13 ; 14 ; 15.

$11\times\frac{1}{4}=2,75$. Comme il faut qu'au moins 2,75 valeurs soient inférieures ou égale à Q_1, il faut que Q_1 soit la 3ème valeur donc $Q_1 = 7$.

$11\times\frac{3}{4}=8,25$. Comme il faut qu'au moins 8,25 valeurs soient inférieures ou égale à Q_3, il faut que Q_3 soit la 9ème valeur donc $Q_3 = 13$.On obtient $\mathbf{Q_1} = 7$, Me =11 et $\mathbf{Q_3} = 13$

- Notes de 10 élèves : 5 ; 6 ; 7 ; 7 ; 9 ; 12 ; 12 ; 13 ; 14 ; 15. $\mathbf{Q_1} = 7$, Me =10,5 et $\mathbf{Q_3} = 13$

(Q_1 est la 3ème valeur et Q_3 la 8ème)

- Notes de 12 élèves : 5 ; 6 ; 7 ; 7 ; 9 ; 11 ; 12 ; 12 ; 13 ; 14 ; 15 ; 16 $\mathbf{Q_1} = 7$, Me =11,5 et $\mathbf{Q_3} = 13$

(Q_1 est la 3ème valeur et Q_3 la 9ème)

Remarque :

L'écart interquartile mesure la dispersion des valeurs autour de la médiane ; plus l'écart est petit, plus les valeurs de la série appartenant à l'intervalle interquartile sont concentrées autour de la médiane.

Contrairement à **l'étendue** (notée *e*) qui mesure l'écart entre la plus grande et la plus petite valeur, l'écart interquartile élimine les valeurs extrêmes qui peuvent être douteuses, cependant il ne tient compte que de 50% de l'effectif.

Remarque :

À la calculatrice, de la même façon que l'on a calculer l'écart-type plus haut, on obtient les quartiles et la médiane en appuyant sur la flèche qui descend.