Statistiques descriptives

I - Quantiles d'une série statistique

Définition *On considère une série statistique d'effectif total N.*

- *La* **médiane** *est une valeur qui partage la série ordonnée en deux séries de même effectif.*
  Si N est impair, $N = 2n+1$, alors la médiane est la $(n+1)^{ème}$ valeur de la série ordonnée.
  Si N est pair, $N = 2n$, alors la médiane est la moyenne de la $n^{ème}$ et de la $(n+1)^{ème}$ valeur.
- *Le* **premier quartile** *Q_1 est une valeur de la série telle que $25\,\%$ des termes de la série lui sont inférieurs.*
  Le **troisième quartile** *Q_3 est une valeur de la série telle que $75\,\%$ des termes de la série lui sont inférieurs.*
  Si N est un multiple de 4, $N = 4n$, alors Q_1 est la $n^{ème}$ valeur et Q_3 la $3n^{ème}$ valeur de la série ordonnée.
  Si N n'est pas un multiple de 4, Q_1 et Q_3 sont les valeurs de rang directement supérieur à $N/4$ et à $3N/4$.
  On appelle **interquartile** *le nombre $Q_3 - Q_1$. C'est une caractéristique de dispersion de la série.*
- *Les* **déciles** *D_1, D_2, D_3, $\dots D_9$ sont des valeurs de la série telles que $10\,\%$, $20\,\%$, $30\,\%$, $\dots$, $90\,\%$ des termes de la série lui sont inférieurs.*

Exercice 1 Soit la série des notes des élèves d'une classe :

Elèves	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Notes	15	10	12	8	10	18	12	8	8	15	10	8	6	18	12	8	12

On ordonne la série :

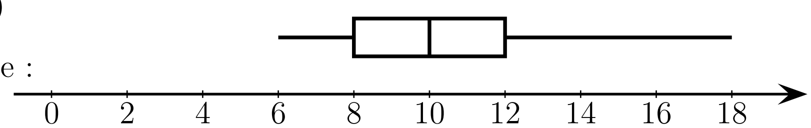
Notes x_i						
Nombre d'élèves n_i						
Effectifs cumulés croissants						

Effectif total de la série : $N = \dots$ Médiane de la série : $M_e = \dots$

Quartiles : $Q_1 = \dots$, $Q_3 = \dots$ Ecart inter-quartile :

Diagrammes en boîte (boites à moustaches)

On représente ces grandeurs sous la forme suivante :



Exercice 2 On compare les températures moyennes (en ° C) de chaque mois de l'année pour deux communes de Haute-Savoie situées à 1000 m d'altitude : Chamonix et La Clusaz.

Mois	1	2	3	4	5	6	7	8	9	10	11	12
Chamonix	1,5	4	7,5	12	15,5	20	23	22	19	14	6,5	2
La Clusaz	2,5	3,5	6	9,5	14	17	20,5	20,0	17	13	7	3,5

Tracer les diagrammes en boîte de ces deux séries.

Exercice 3 Trois classes ont passé un devoir commun. Les notes sont les suivantes :

— classe A : 5 ; 8 ; 6 ; 3 ; 11 ; 4 ; 8 ; 15 ; 8 ; 11 ; 14 ; 12 ; 10 ; 10 ; 1 ; 12 ; 14 ; 14 ; 5 ; 6 ; 10 ; 14 ; 10 ; 8 ; 9 ; 0 ; 8 ; 10 ; 6 ; 16 ; 12

— classe B : 13 ; 2 ; 13 ; 13 ; 16 ; 16 ; 18 ; 11 ; 15 ; 10 ; 15 ; 11 ; 15 ; 11 ; 19 ; 12 ; 6 ; 10 ; 17 ; 11 ; 12 ; 5 ; 16 ; 10 ; 17 ; 18 ; 2 ; 16 ; 4 ; 9 ; 14 ; 12

— calsse C : 9 ; 7 ; 12 ; 8 ; 9 ; 8 ; 11 ; 12 ; 10 ; 12 ; 13 ; 9 ; 15 ; 9 ; 14 ; 15 ; 9 ; 14 ; 13 ; 16 ; 5 ; 17 ; 13

Comparer les trois classes.

II - Moyenne, variance et écart-type

On considère une série statistique générale :

Valeurs du caractère	x_1	x_2	x_3	...	x_p
Effectifs	n_1	n_2	n_3	...	n_p

d'effectif total : $N = n_1 + n_2 + \cdots + n_p$.

Définition *La moyenne de la série est :* $\overline{x} = \dfrac{n_1x_1 + n_2x_2 + \cdots + n_px_p}{N} = \dfrac{1}{N}\displaystyle\sum_{i=1}^{p} n_ix_i$

Exercice 4 Déterminer la moyenne des séries suivantes :

S_1 : 1 ; 8 ; 10 ; 10 ; 12 ; 19 $\qquad \overline{x} = \ldots.$

S_2 : 9 ; 9,5 ; 10 ; 10,5 ; 11 $\qquad \overline{x} = \ldots.$

S_3 : 10 ; 10 ; 10 ; 10 ; 10 ; $\qquad \overline{x} = \ldots.$

S_4 : 3 ; 7 ; 12 ; 16 ; 18 ; 8 ; 9,5 ; 12,5 ; 6 ; 8 $\qquad \overline{x} = \ldots.$

Définition

- *La variance de la série est la moyenne des carrés des écarts à la moyenne :*

$$V = \frac{n_1(x_1-\overline{x})^2 + n_2(x_2-\overline{x})^2 + \cdots + n_p(x_p-\overline{x})^2}{N} = \frac{1}{N}\sum_{i=1}^{p} n_i(x_i-\overline{x})^2$$

- *L'écart-type d'une série est la racine carrée de la variance :* $\sigma = \sqrt{V}$.

Propriété *La variance est la moyenne des carrés des valeurs de la série moins le carré de la moyenne de la série :* $V = \overline{x^2} - \overline{x}^2$

Démonstration :

$$V = \frac{1}{N}\sum_{i=1}^{p} n_i(x_i-\overline{x})^2 = \frac{1}{N}\sum_{i=1}^{p} n_i\Big(\quad \ldots \quad \ldots \quad \Big)$$
$$= \ldots$$

Exercice 5 Le tableau suivant donne les tailles de 35 élèves d'une classe.

taille(cm)	145	146	151	152	155	160	165	170	172	176	180	186	188	190	193
effectif	1	1	1	2	3	3	5	2	6	3	3	2	1	1	1

Calculer la moyenne et l'écart-type de cette série.

Exercice 6

1. On considère l'algorithme :

```
Demander N
Affecter 0 à S
Pour i de 1 à N
   Demander x
   Affecter S+x à S
Fin
Affecter S/N à M
Afficher M
```

   a) Qu'affiche cet algorithme si on entre 4 pour N, puis successivement les valeurs 8, 15, 7, 10.
   b) Que fait cet algorithme ?
   c) Programmer cet algorithme (python, calculatrice, ...)

2. Modifier l'algorithme/programme précédent pour que le calcul de la moyenne puisse prendre en compte des coefficients (à demander avec chaque valeur de x).
3. En utilisant l'algorithme/programme précédent, écrire un algorithme/programme qui calcule la variance et l'écart type de la série de valeurs entrée.