

DOSSIER 4 - Valeurs de position

1 Savoirs

1.1 Variable discrète

Pour des caractères quantitatifs, on peut décrire la série statistique au moyen de nombres. La notion de moyenne est connue. On va la (re)définir et définir d’autres valeurs de position. Pour cela, partons de l’exemple des températures. Son diagramme en bâtonnets est donné à la figure 1 et les tableaux bruts et de calculs aux tables 1 et 2.

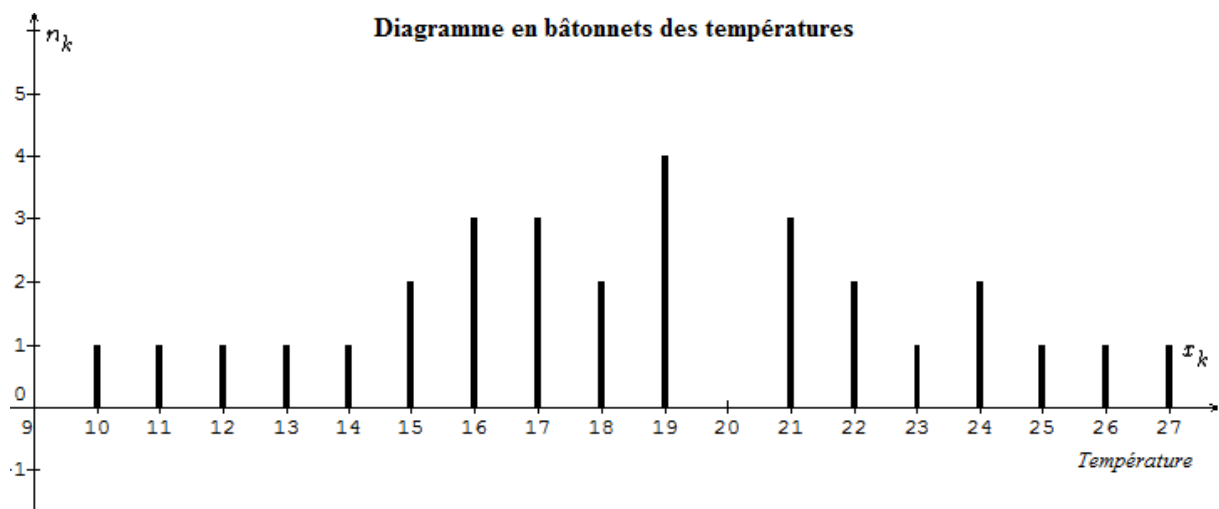


FIGURE 1 – Diagramme en bâtonnets des températures

12	10	11	13	15	16	16	17	19	18
19	17	16	15	14	17	19	18	19	21
22	21	21	23	22	24	25	27	26	24

TABLE 1 – Températures relevées en milieu de journée

- Le **mode** de la série vaut 19. C’est la modalité dont l’effectif est le plus élevé. On peut aussi dire que c’est la valeur du caractère auquel correspond l’effectif le plus élevé.
- La **moyenne arithmétique** des nombres $x_1, x_2, \dots, x_n$, notée m , est le nombre

$$m = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i.$$

On peut aussi calculer la moyenne à partir des données du tableau recensé. Si n_k est l’effectif de la modalité x_k et n est l’effectif total, alors

$$m = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{n} = \frac{1}{n} \sum_{i=1}^p n_ix_i.$$

Modalités x_k	Effectifs n_k	Fréquences (%) $f_k = \frac{n_k}{n}$	Effectifs cumulés N_k	Fréquences cumulées F_k
10	1	$\frac{1}{30} \cdot 100 = 3,33$	1	3,33
11	1	3,33	2	6,67
12	1	3,33	3	10
13	1	3,33	4	13,33
14	1	3,33	5	16,67
15	2	6,67	7	23,33
16	3	10	10	33,33
17	3	10	13	43,33
18	2	6,67	15	50
19	4	13,33	19	63,33
21	3	10	22	73,33
22	2	6,67	24	80
23	1	3,33	25	83,33
24	2	6,67	27	90
25	1	3,33	28	93,33
26	1	3,33	29	96,67
27	1	3,33	30	100
Total	n=30	100		

TABLE 2 – Tableau de calcul des températures

Pour l'exemple 2, cela donne :

· Sur base du tableau brut (table 1) : $m = \frac{12+10+11+\dots+26+24}{30} = \frac{557}{30} = 18,57$

· Sur base du tableau de calcul (table 2) : $m = \frac{1 \cdot 10 + 1 \cdot 11 + \dots + 2 \cdot 15 + 3 \cdot 16 + \dots + 1 \cdot 27}{30} = \frac{557}{30} = 18,57$

- La **médiane**, M , d'une série statistique à caractère quantitatif est la valeur du caractère telle que la **moitié de l'effectif total** lui soit inférieur ou égale et l'autre moitié lui soit supérieure ou égale. Elle se trouve au centre des observations.

Si n est impair, alors M est l'élément central de la série et si n est pair, alors M est la moyenne des deux éléments centraux de la série.

À partir du diagramme des fréquences cumulées, M est l'abscisse du point qui a 50% comme ordonnée. Pour le diagramme des effectifs cumulés, l'ordonnée à prendre en compte est la moitié de l'effectif total.

Sur base de son diagramme des effectifs cumulés (fig. 2), la médiane de la série de l'exemple des températures vaut 18,5.

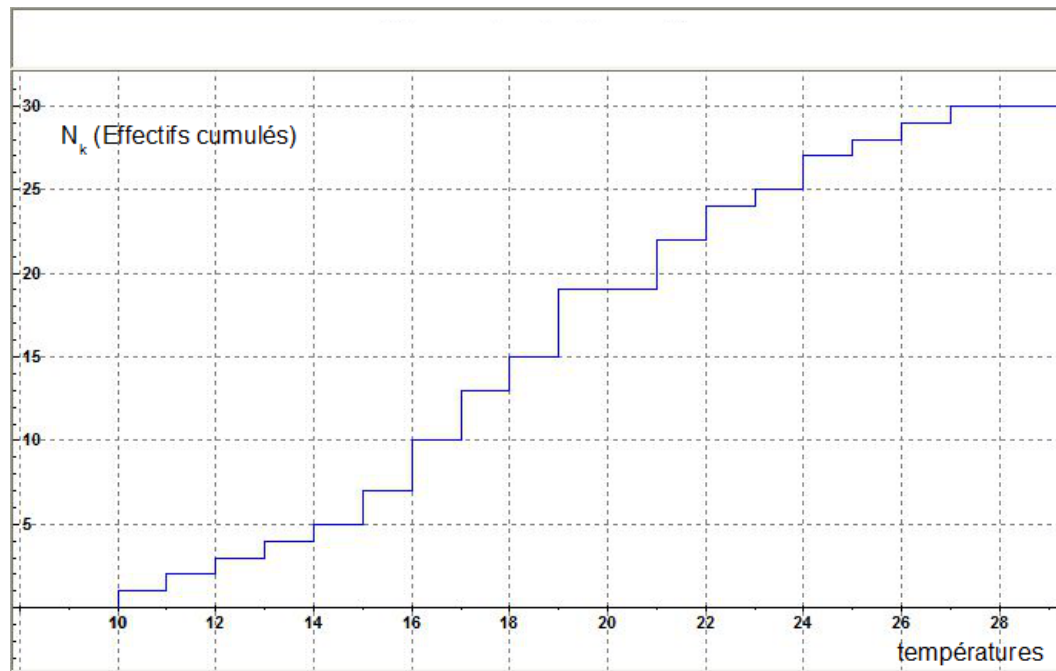


FIGURE 2 – Diagramme des effectifs cumulés des températures

- Les quartiles partagent la population en quatre groupes de même effectif.
Sur le diagramme des fréquences cumulées,
 - le **premier quartile**, noté Q_1 , correspond à l'abscisse du point dont l'ordonnée est 0,25 ;
 - le **deuxième quartile**, noté Q_2 , correspond à l'abscisse du point dont l'ordonnée est 0,5, c'est la médiane ;
 - le **troisième quartile**, noté Q_3 , correspond à l'abscisse du point dont l'ordonnée est 0,75.

Si on part du diagramme des effectifs cumulés, il faut prendre les abscisses des points d'ordonnées $\frac{n}{4}$, $\frac{n}{2}$ et $\frac{3n}{4}$, respectivement. Si l'ordonnée correspond à un segment horizontal, on prendra la moyenne des extrémités du segment.

Sur base de son diagramme des effectifs cumulés (fig. 2), les quartiles de la série de l'exemple des températures valent : $Q_1 = 16$, $Q_2 = M = 18,5$ et $Q_3 = 22$

1.2 Variable continue

Nous repartons de l'exemple de la taille des pins du dossier 3 dont on redonne l'histogramme et le polygone des fréquences cumulées.

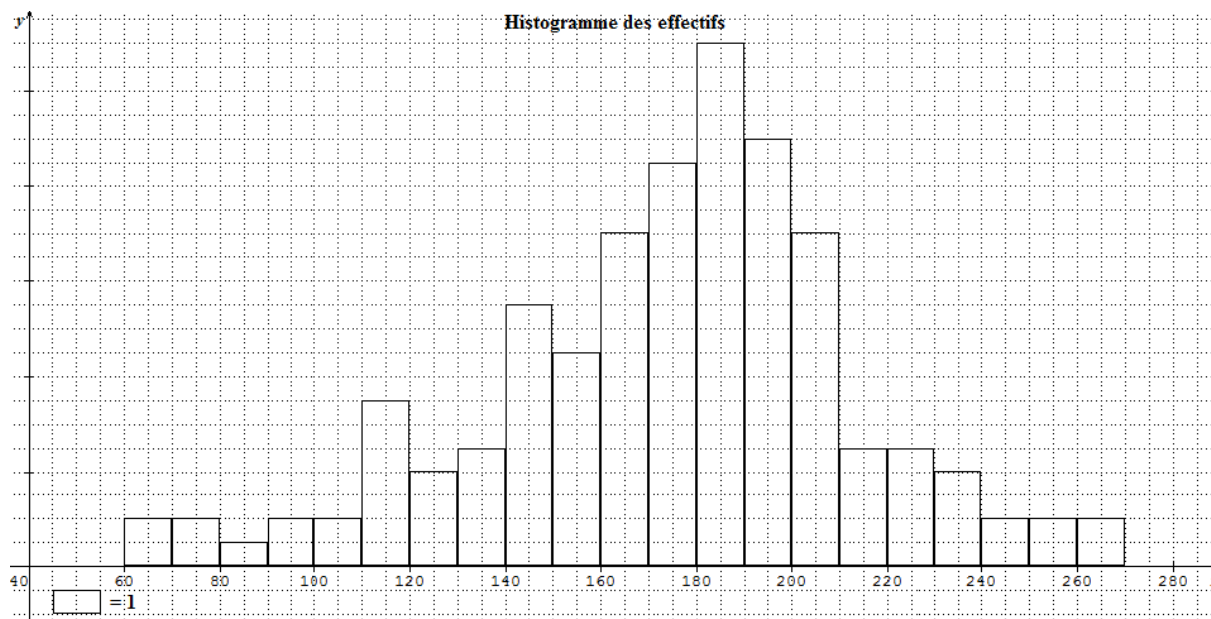


FIGURE 3 – Histogramme des effectifs de la taille de pins

- Dans le cas continu, la **classe modale** est la classe dont l'effectif est le plus élevé. Le **mode** est alors donné par le centre de cette classe. Pour notre exemple, il vaut :
- Pour calculer la **moyenne** quand les données ont été regroupée en classes, il faut considérer les **centres** c_k des classes x_k .

$$m = \frac{1}{n} \sum_{i=1}^p n_i c_i.$$

Pour notre exemple, cela donne $m = \frac{2.65+2.75+1.85+2.95+...+2.165}{150} = \frac{26140}{150} = 174,3$

- Pour obtenir la **médiane** et les **quartiles** Q_1 , M et Q_3 , on se réfère au polygone des fréquences (ou effectifs) cumulées comme dans le cas discret.

Sur base du polygone des fréquences cumulées (fig.4), on obtient :

$Q_1 = 152$, $M = 179$ et $Q_3 = 198$

- La **boîte à moustaches** (qu'on appelle aussi diagramme en boîte ou boîte de Tukey) résume les paramètres de positions en y ajoutant les valeurs extrêmes. La boîte est composée d'un rectangle qui va du premier au troisième quartile et qui est coupé par la médiane. Les moustaches sont les segments qui relient cette boîte aux valeurs extrêmes.

Dans un traitement continu, les valeurs extrêmes sont les extrémités de la première et de la dernière classe. Nous avons appris à trouver les valeurs des quartiles sur le polygone des fréquences ou des effectifs cumulés. C'est donc à partir de ces graphes que nous pourrions construire une boîte à moustaches.

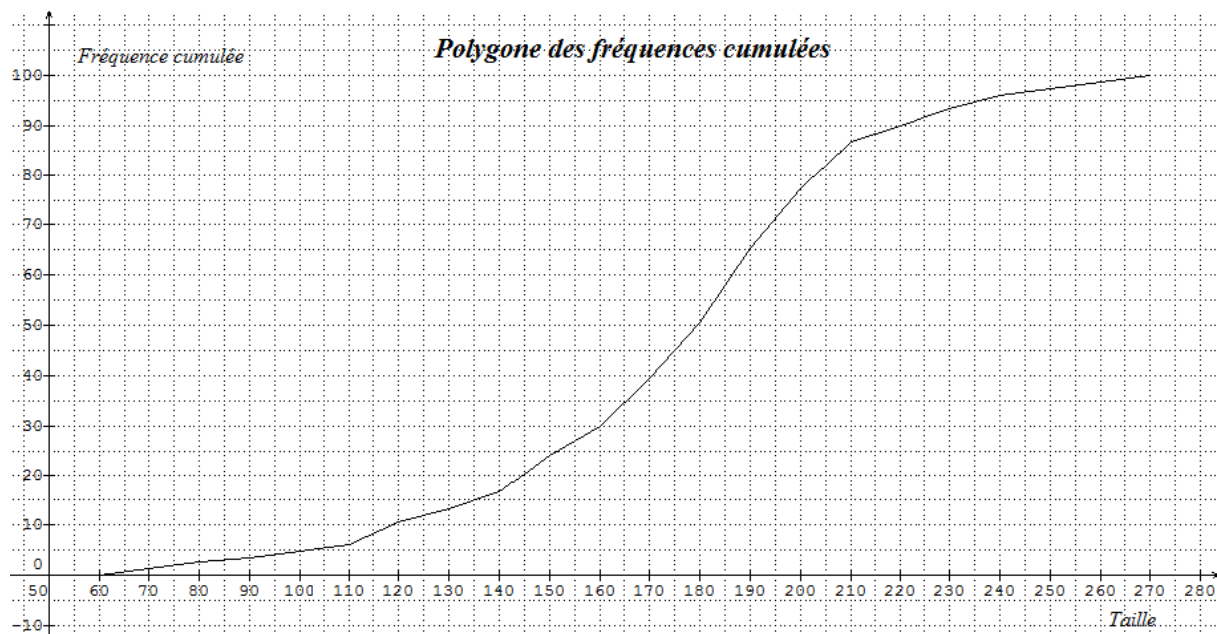


FIGURE 4 – Polygone des fréquences cumulées de la taille de pins

Complétons la figure 5 pour bien montrer comment passer du polygone des effectifs cumulés à la boîte à moustaches.

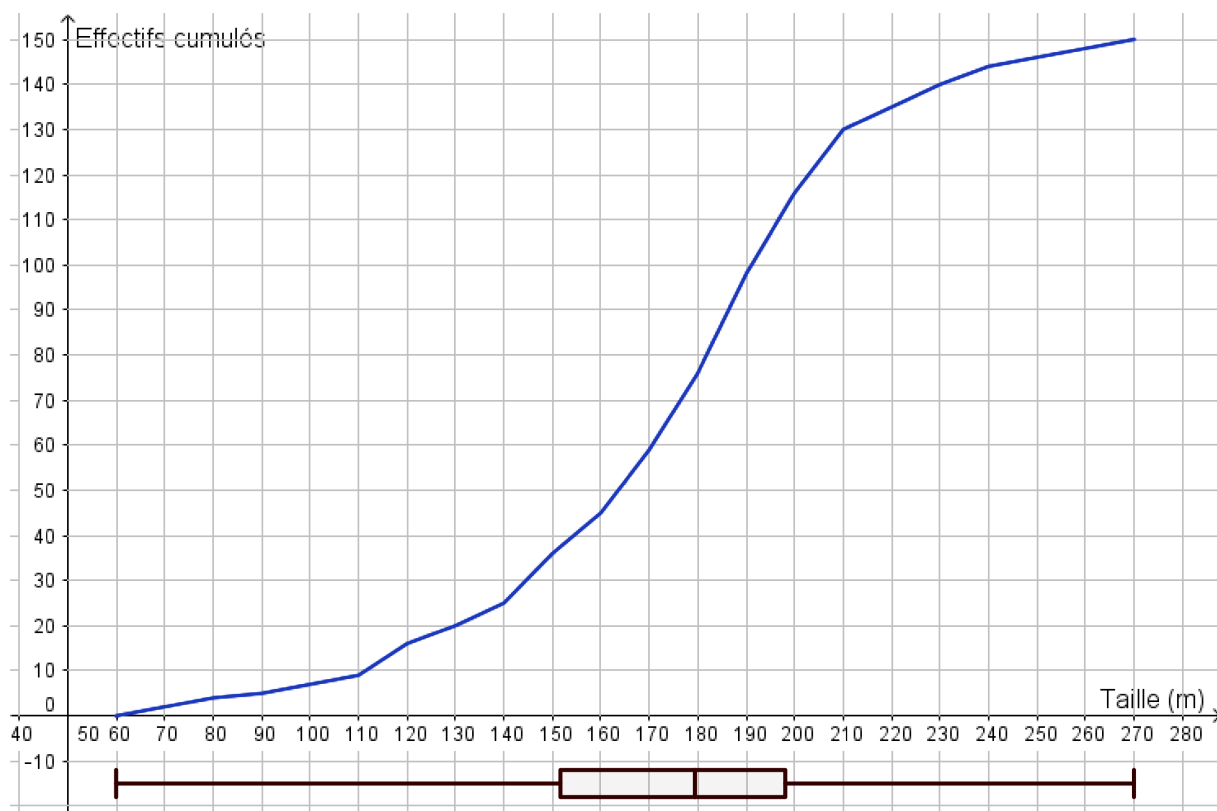


FIGURE 5 – La boîte à moustaches de la taille de pins est obtenue à partir du polygone des effectifs cumulés.

2 Applications

1. Une poule pond huit œufs.

Voici les poids en grammes (g) des œufs : 60 g, 56 g, 61 g, 68 g, 51 g, 53 g, 69 g, 54 g.

Détermine en indiquant les calculs :

- a) le mode, b) la moyenne, c) la médiane.

- ## 2. Note en Math

Voici les notes (sur 100 points) obtenues aux examens par 30 élèves.

75	84	71	85	75	94	77	69	74	93
48	68	93	59	87	74	62	95	35	71
63	72	60	75	78	95	78	85	76	76

Déterminer le tableau de valeur qui permettra par la suite de

- Déterminer la médiane
- Préciser quel est l'écart, en terme de points, entre le meilleur résultat et le moins bon ?
- Déterminer la note en dessous de laquelle les élèves pourront participer au rattrapage sachant que seuls les 10 plus faibles pourront y participer.
- Déterminer la note au dessus de la quelle les élèves pourront participer aux olympiades sachant que les 5 meilleurs seront sélectionnés.

- ### 3. Salaire moyen

Dans une entreprise familiale, la direction est assurée par le directeur, son frère et des parents proches. On annonce un salaire moyen de 3000 € par mois. Pourtant, un nouvel ouvrier ne comprend pas bien. Lui et ses camarades ne gagnent que 1000 € par mois. Comment est-ce possible ?

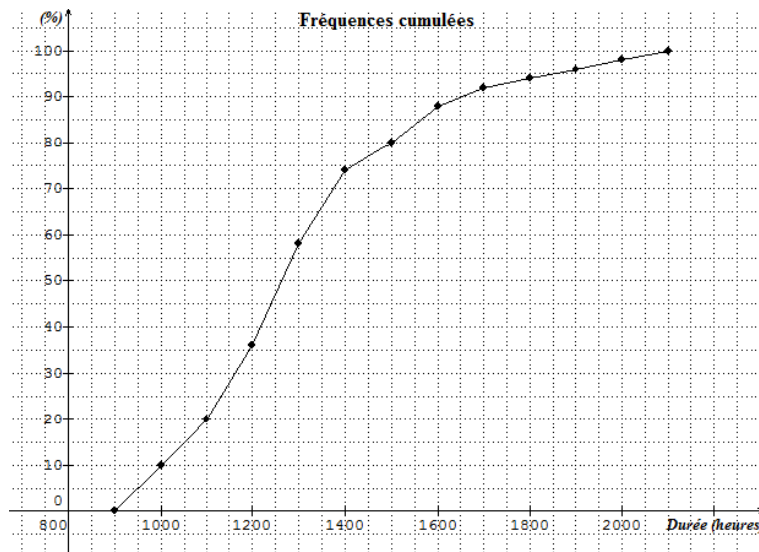
Voici un tableau reprenant les différents postes avec les salaires correspondants.

Poste	Directeur	Frère	Parents	Contremaîtres	Ouvriers
Nombre	1	1	6	5	10
Montant (€)	24000	10000	2500	2000	1000

- Vérifie que la moyenne vaut bien 3000 €.
- Quelle est le mode ? Justifie.
- Quelle est la médiane ? Justifie.
- Compare les trois valeurs. D'où vient le fait que l'ouvrier ne comprenne pas bien la situation ?

4. *Durée de vie de piles*

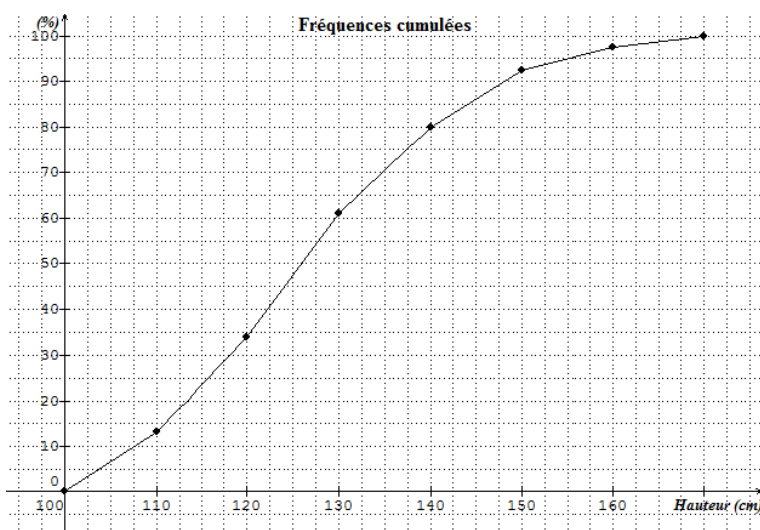
On a étudié la durée de vie de piles et on a établi le polygone des fréquences cumulées.



- Détermine la médiane et les quartiles.
- Dessine la boîte à moustache correspondant à ces données.
- Complète les phrases :
 - Il y a ... % des piles qui durent au maximum 1000 heures.
 - Il y a ... % des piles qui durent plus de 1200 heures.
 - Il y a ... % des piles qui durent entre 1200 et 1400 heures.

5. *Hauteurs de buissons*

On a étudié la hauteur des buissons dans une zone d'habitation et on a établi le polygone des fréquences cumulées.



- Dessine la boîte à moustache correspondant à ces données en précisant quelles sont les valeurs que tu as utilisées
- Quel est le mode de cette série ?
- Pourrait-on réaliser un histogramme des effectifs et/ou des fréquences à partir du graphe donné ? Si oui, comment ?

6. *Budget vacances*

Un agence de voyage de la région du centre à réalisé une étude statistique sur les budgets consacrés aux vacances d'été (par personne) dans la région du centre afin de peaufiner son catalogue d'offres.

Budget par personne (€)	[300; 600[	[600; 900[	[900; 1200[	[1200; 1500[	[1500; 1800[
Nombre de réponses	241	531	352	124	52

A l'aide d'un polygone des fréquences cumulées, réponds aux questions suivantes :

- En moyenne quel est le budget consacré aux vacances d'été dans la région du centre ?
- L'agence souhaite pouvoir répondre principalement aux demandes de 75% de la population. Quel est le montant maximum des voyages qu'elle doit proposer ?
- Sur la première page du catalogue, elle souhaite mettre en évidence un voyage qui pourrait convenir à 50% de la population. Pour une famille de 4 personnes, quel sera le prix de celui-ci ?