

Vocabulaire de la statistique descriptive

avec le centre d'une classe est $c_i = \frac{b_i - b_{i-1}}{2}$.

et l'amplitude d'une classe est $a_i = b_i - b_{i-1}$

Exemple 1.5.3

La répartition de 100 ménages selon leurs dépenses de consommation mensuelles exprimées en milliers dinars se présente comme suit :

| <i>Classes de dépenses</i> | <i>Nombre de ménages</i> |
|----------------------------|--------------------------|
| [20, 40[| 15 |
| [40, 60[| 20 |
| [60, 100[| 20 |
| [100, 200[| 45 |

et le tableau associé est :

| <i>Classes $[b_{i-1}, b_i[$</i> | <i>Centres c_i</i> | n_i | N_i | f_i | F_i |
|--|---------------------------------|------------|-------|-------------|-------|
| [20, 40[| 30 | 15 | 15 | 0,15 | 0,15 |
| [40, 60[| 50 | 20 | 35 | 0,20 | 0,35 |
| [60, 100[| 80 | 20 | 55 | 0,20 | 0,55 |
| [100, 200[| 150 | 45 | 100 | 0,45 | 1,00 |
| Total | | 100 | | 1,00 | |

Exemple 1.5.4 : (le calcul de F'_i et N'_i)

On calcul les effectifs cumulés décroissants et les fréquences cumulées croissantes comme apparait dans le tableau

| <i>Classes $[b_{i-1}, b_i[$</i> | <i>Centres c_i</i> | n_i | N'_i | f_i | F'_i |
|--|---------------------------------|------------|--------|-------------|--------|
| [20, 40[| 30 | 15 | 100 | 0,15 | 1,00 |
| [40, 60[| 50 | 20 | 85 | 0,20 | 0,85 |
| [60, 100[| 80 | 20 | 65 | 0,20 | 0,65 |
| [100, 200[| 150 | 45 | 45 | 0,45 | 0,45 |
| Total | | 100 | | 1,00 | |

Remarque 1.5.5 L'ensemble des couples

1. $\{(x_i, n_i)\}$ ou encore $\{(x_i, f_i)\}$ si la variable est discrète.
2. $\{([b_{i-1}, b_i[, n_i)\}$, ou $\{([b_{i-1}, b_i[, f_i)\}$ si la variable est continue.

Est appelé série statistique de la variable.

1.6 Séries statistiques et leurs représentations graphiques

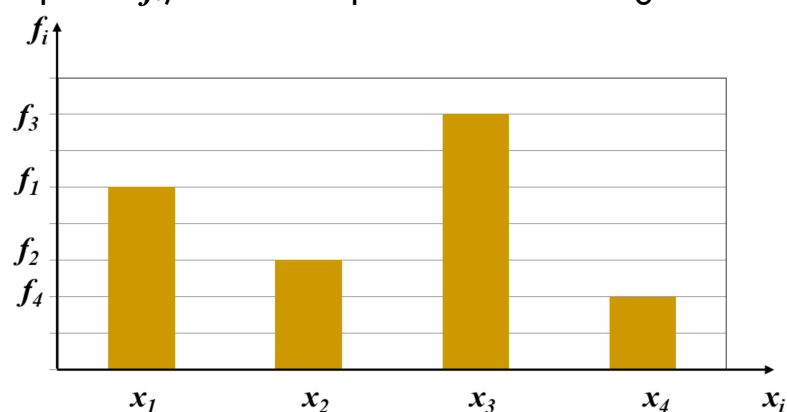
Un tableau statistique contient toute l'information prélevée sur un échantillon mais il est très utile de présenter ces informations par un graphique, afin d'avoir un aperçu de l'évolution de la variable. On utilise des différents types de représentations graphiques suivant la nature de la variable étudiée.

1.6.1 Cas qualitatif

- **Diagramme à bandes**

C'est un repère cartésien tel que :

à chaque modalité x_i on associe un rectangle de base constante dont la hauteur est l'effectif n_i (la fréquence f_i) comme est présenté dans la figure suivante

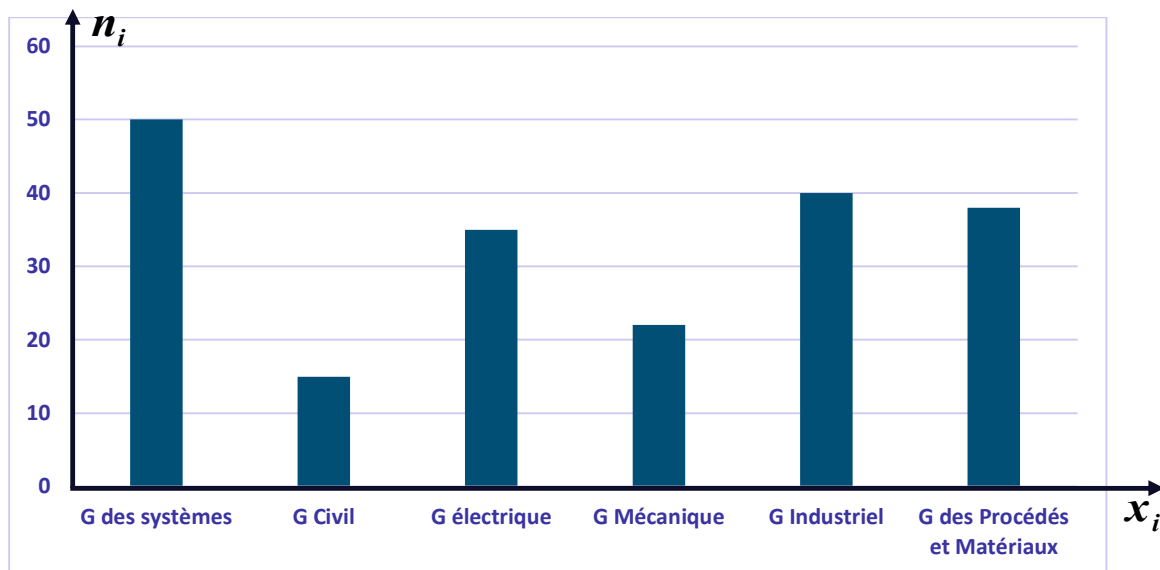


Remarque 1.6.1 Pour l'axe des effectifs (fréquences), on choisit une échelle arithmétique.

Exemple 1.6.1 D'après une étude faite à l'ENPO-Oran, la répartition de 200 étudiants de deuxième année des CPST admis en troisième année, sur les spécialités préférées a fourni les résultats suivants

| <i>Spécialité préférée x_i</i> | <i>Nombre d'élèves n_i</i> |
|---|---|
| <i>Génie des systèmes</i> | 50 |
| <i>Génie Civil</i> | 15 |
| <i>Génie électrique</i> | 35 |
| <i>Génie Mécanique</i> | 22 |
| <i>Génie Industriel</i> | 40 |
| <i>Génie des Procédés et Matériaux</i> | 38 |
| <i>Total</i> | 200 |

Ces résultats peuvent être traduits par le diagramme à bandes suivant

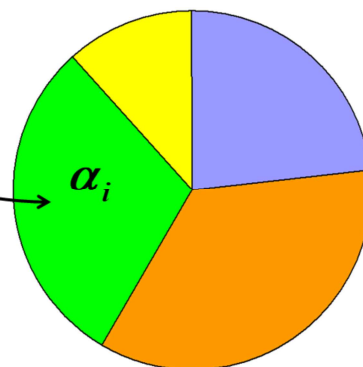


- **Diagramme à secteurs (circulaires)**

C'est un graphique où, les modalités sont représentées par des portions de disque proportionnelles à leurs effectifs, ou à leurs fréquences.

En effet, pour une modalité x_i , d'effectif n_i , l'angle au centre α_i correspondant est donné (en degré) par :

$$\alpha_i = f_i \times 360^\circ = \frac{n_i}{N} \times 360^\circ$$



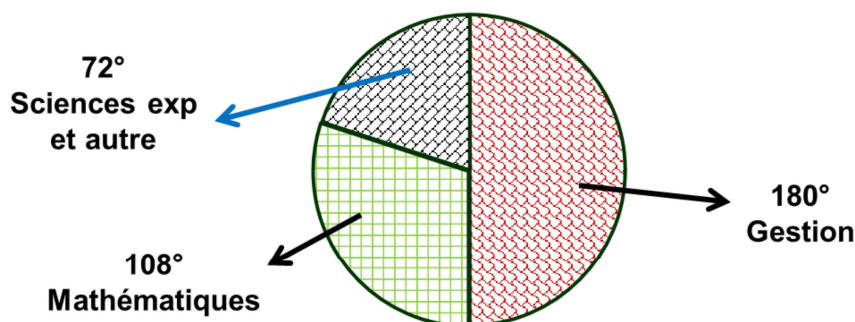
Remarque 1.6.2 Les diagrammes à bandes et circulaire peuvent être utilisés dans le cas quantitatif.

Exemple 1.6.2

D'après une étude faite à l'école de commerce d'Oran, la répartition de 50 étudiants selon la branche du bac est reportée dans le tableau suivant :

| <i>Section du bac x_i</i> | <i>Effectifs n_i</i> | <i>f_i</i> | <i>α_i</i> |
|--|-----------------------------------|-------------------------|------------------------------|
| <i>Gestion</i> | 25 | 0,50 | 180° |
| <i>Mathématiques</i> | 15 | 0,30 | 108° |
| <i>Sciences exp et autres</i> | 10 | 0,20 | 72° |
| <i>Total</i> | 50 | 1,00 | 360° |

Ces données peuvent être traduits par le diagramme à secteurs suivant



1.6.2 Cas quantitatif discret (Diagrammes en bâtons)

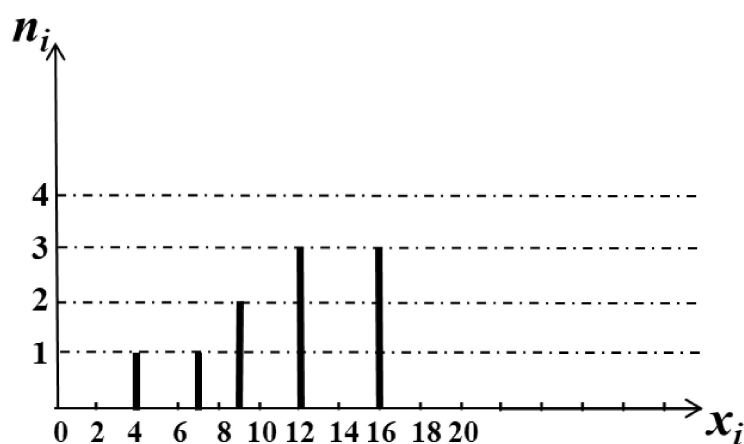
C'est un repère cartésien tel que les valeurs sont placées en abscisse, les effectifs (ou fréquences) en ordonnée, et à chaque point $(x_i, 0)$ on associe un segment vertical dont la longueur est l'effectif n_i (la fréquence f_i), voir les deux exemples suivants.

Exemples 1.6.3

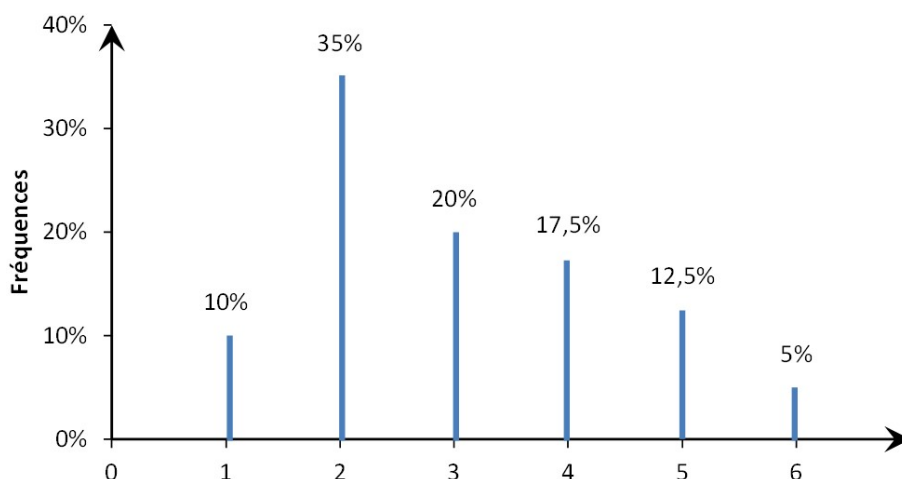
La répartition des notes d'un contrôle noté sur 20 pour une classe de 10 élèves est reportée dans le tableau suivant :

| x_i | <i>Effectifs n_i</i> |
|--------------|-----------------------------------|
| 4 | 1 |
| 7 | 1 |
| 9 | 2 |
| 12 | 3 |
| 16 | 3 |
| <i>Total</i> | 10 |

donne



La répartition d'une population de familles selon le nombre d'enfant de chaque famille est représentée sous le diagramme en bâtons suivant

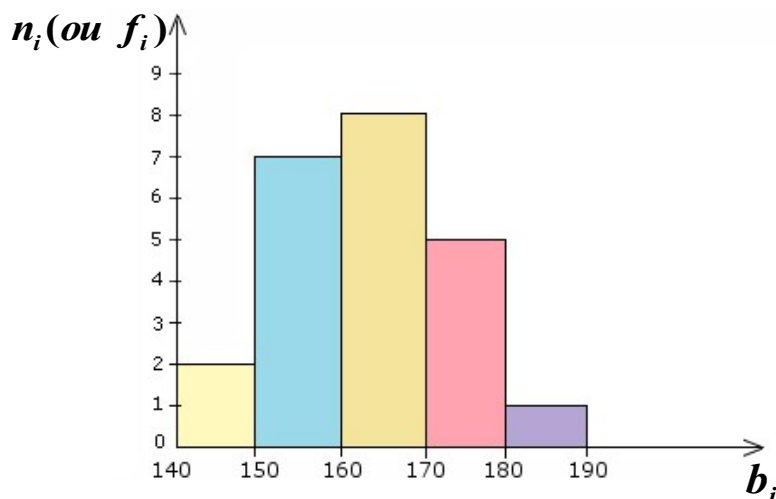


A partir de ce diagramme on peut construire le tableau des données suivant

| x_i | Fréquences en (%) f_i |
|--------------|-------------------------|
| 1 | 10 |
| 2 | 35 |
| 3 | 20 |
| 4 | 17,5 |
| 5 | 12,5 |
| 6 | 5 |
| Total | 100 % |

1.6.3 Cas quantitatif continu (Histogramme)

On représente une série statistique continue par un histogramme. Il s'agit d'une figure obtenue sur un repère cartésien en représentant pour chaque classe $[b_{i-1}, b_i[$ un rectangle de surface S_i proportionnelle à l'effectif n_i ou à la fréquence f_i . Les rectangles de l'histogramme sont voisins, comme est présenté dans la figure suivante



• Principe de construction de l'histogramme

Il y'a deux (02) cas

1^{er} cas : Cas d'amplitudes égales

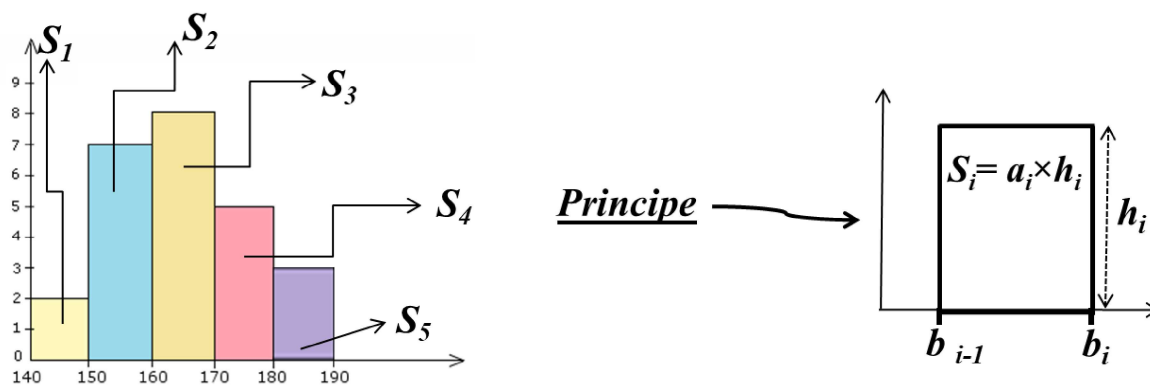
Si les classes sont de même amplitude a_i (ie $a_i = a_j$), on place en ordonnée les effectifs n_i (ou les fréquences f_i).

Exemple 1.6.4

La répartition de 20 personnes selon leurs tailles est donnée dans le tableau suivant :

| $[b_{i-1}, b_i[$ | n_i | a_i | f_i |
|------------------|-------|-------|-------|
| $[140, 150[$ | 2 | 10 | 0,08 |
| $[150, 160[$ | 7 | 10 | 0,28 |
| $[160, 170[$ | 8 | 10 | 0,32 |
| $[170, 180[$ | 5 | 10 | 0,20 |
| $[180, 190[$ | 3 | 10 | 0,12 |

Ces données peuvent être représentées par l'histogramme suivant



2^{eme} cas : Cas d'amplitudes inégales

Si les classes sont d'amplitude inégales a_i (ie $a_i \neq a_j$), on définit

1. La densité d'une classe par : $d_i = \frac{n_i}{a_i}$

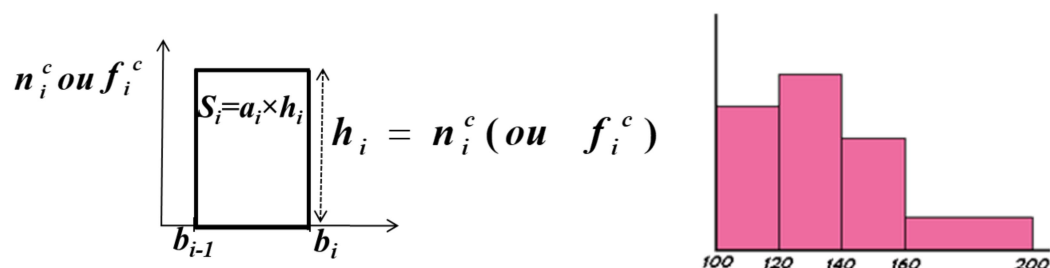
et on pose

$$h_i = \frac{n_i}{a_i} \times a^* = d_i \times a^* = n_i^c$$

2. avec a^* est appelée **amplitude de référence**. Elle est choisie arbitrairement de manière à faciliter la représentation graphique (valeurs sur l'axe des ordonnées).

3. h_i est dans ce cas est appelée **effectif corrigé** qu'on note n_i^c .

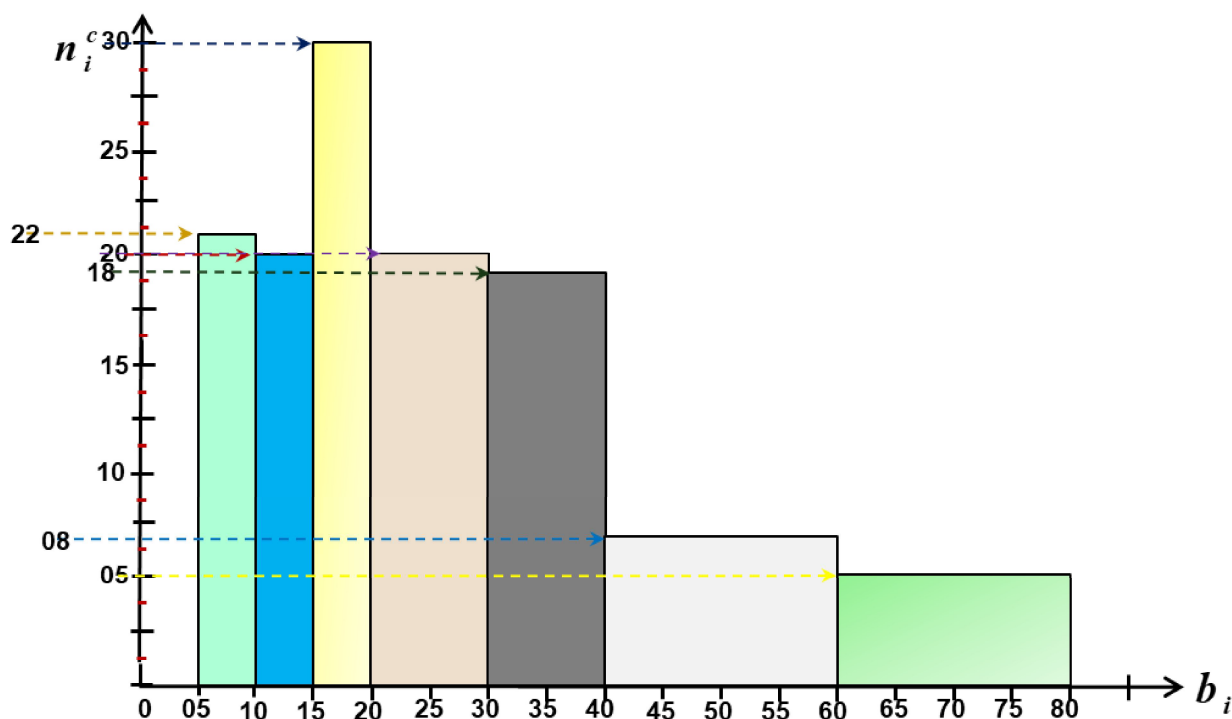
Ainsi les rectangles S_i seront de la manière suivante :



Exemple 1.6.5 La répartition de 100 individus par classes d'âges est donnée par le tableau suivant :

| Classe | n_i | a_i | $d_i = \frac{n_i}{a_i}$ | $n_i^c = d_i \times a^*$ | f_i | $f_i^c = \frac{n_i^c}{N} = \frac{f_i}{a_i} \times a^*$ |
|--------------|------------|-------|-------------------------|--------------------------|-------------|--|
| [05 , 10[| 11 | 5 | 2,2 | 22 | 0,11 | 0,22 |
| [10 , 15[| 10 | 5 | 2,0 | 20 | 0,10 | 0,20 |
| [15 , 20[| 15 | 5 | 3,0 | 30 | 0,30 | 0,30 |
| [20 , 30[| 20 | 10 | 2,0 | 20 | 0,20 | 0,20 |
| [30 , 40[| 18 | 10 | 1,8 | 18 | 0,18 | 0,18 |
| [40 , 60[| 16 | 20 | 0,8 | 08 | 0,16 | 0,08 |
| [60 , 80[| 10 | 20 | 0,5 | 05 | 0,10 | 0,05 |
| Total | 100 | | | | 1,00 | |

Si on prend l'amplitude de référence $a^* = 10$ on obtient l'histogramme des effectifs suivant :



Remarque 1.6.3

- Il existe deux types d'histogramme (histogramme des effectifs et histogramme des fréquences).
- On associe à un histogramme des fréquences un polygone des fréquences qui se définit comme suit :

- **Polygone des fréquences**

Il s'agit d'une ligne brisée reliant :

- les milieux des sommets des rectangles de l'histogramme.
- La fermeture se fait par deux points sur l'axe des abscisses situés respectivement à un demi-intervalle de La borne inférieure de la première classe et de la borne supérieure de la dernière classe.

Pour bien voir un cas d'un histogramme avec son polygone des fréquences faisant l'exercice suivant.

Exercice d'application

Une association de course à pieds a une équipe féminine. La liste suivante est composée des prénoms d'athlètes suivis entre parenthèses du derniers temps aux 10 Km.

Aicha(51), Ahlem(49), Amel(50), Badra(58), Bouchra(55), Dalia(64), Fadia(60), Fahima(61), Fatiha(46), Fatima(56), Fouzia(50), Hajera(42), Houria(54), Ikram(48), Ilham(45), Imane(57), Jamila(59), Khadija(54), Lamia(54), Leila(46), Meriem(46),

Nabila(41), Samia(39), Samira(37), Wafaa(50), Yamina(47), Yasmine(50), Zahira(44), Zakia(51), Zoulikha(59).

Le responsable de l'association décide de créer dans un ordre croissant cinq (05) équipes (classes) de niveau d'athlètes équivalents telles que :

la 1^{ère} équipe contient 3 athlètes, la 2^{ème} équipe contient 3 athlètes, la 3^{ème} équipe contient 6 athlètes, la 4^{ème} équipe contient 9 athlètes, et la 5^{ème} équipe contient 9 athlètes.

1. Constituer les équipes. (Faire un tableau donnant les temps minimums et maximums pour chacune des équipes).
2. Donner une représentation graphique des fréquences sous forme d'un histogramme (l'amplitude de référence $a^* = 1000$).
3. Dessiner le polygone des fréquences.

Solution de l'exercice d'application

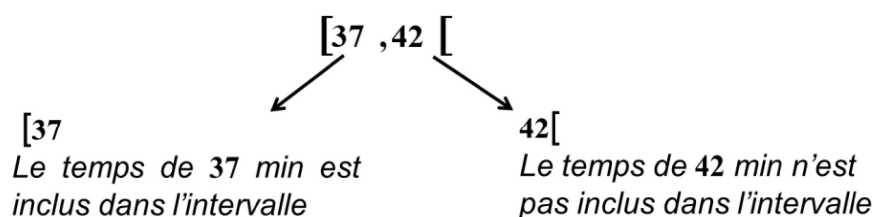
1. **Constitution des équipes** : La liste des athlètes est ;

Aicha(51), Ahlem(49), Amel(50), Badra(58), Bouchra(55), Dalia(64), Fadia(60), Fahima(61), Fatiha (46), Fatima(56), Fouzia(50), Hajera(42), Houria(54), Ikram(48), Ilham(45), Imane(57), Jamila(59), Khadija(54), Lamia(54), Leila(46), Meriem(46), Nabila(41), Samia(39), Samira(37), Wafaa(50), Yamina(47), Yasmine(50), Zahira(44), Zakia(51), Zoulikha(59).

Et on va ranger par ordre croissant des temps : Les 3 sportives dont le temps est le plus petit (les meilleures) constitueront l'équipe 1 ;

| Equipe 1 | Equipe 2 | Equipe 3 | Equipe 4 | Equipe 5 |
|--------------------|--------------------|--------------------|---------------------|----------------------|
| Samira (37) | Hajera (42) | Leila (46) | Amel (50) | Bouchra (55) |
| Samia (39) | Zahira (44) | Fatiha (46) | Wafaa (50) | Fatima (56) |
| Nabila (41) | Ilham (45) | Meriem (46) | Yasmine (50) | Imane (57) |
| | | Yamina (47) | Fouzia (50) | Badra (58) |
| | | Ikram (48) | Aicha (51) | Jamila (59) |
| | | Ahlem (49) | Zakia (51) | Zoulikha (59) |
| | | | Khadija (54) | Fadia (60) |
| | | | Houria (54) | Fahima (61) |
| | | | Lamia (54) | Dalia (64) |

On a créé ainsi des « classes ». On les écrit sous forme d'intervalle, par exemple l'intervalle de temps de l'équipe 1 est :



On peut ainsi construire un nouveau tableau :

Vocabulaire de la statistique descriptive

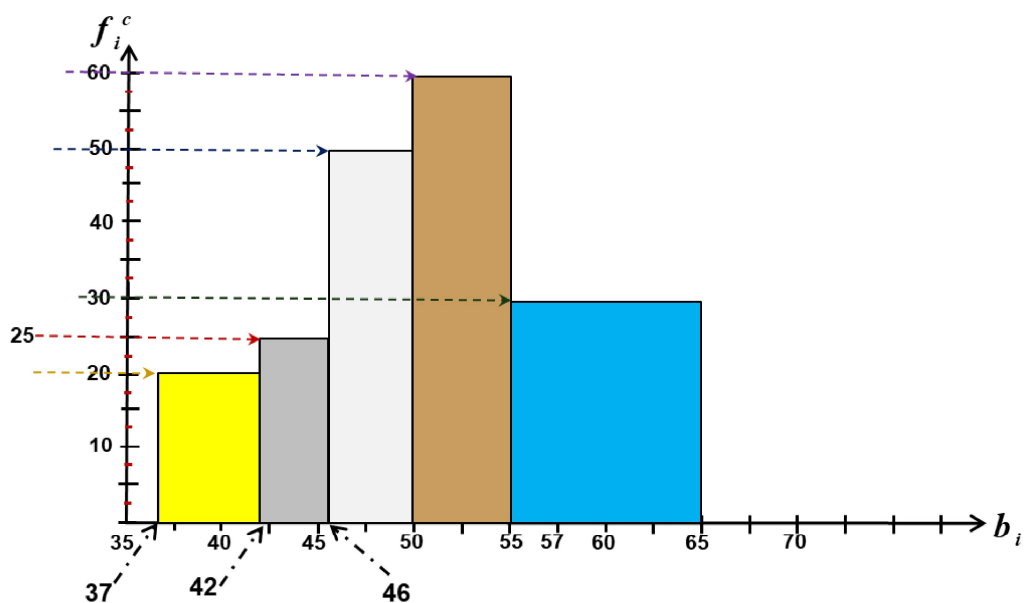
| $[b_{i-1}, b_i[$ | n_i |
|------------------|-------|
| [37, 42[| 3 |
| [42 , 46[| 3 |
| [46 , 50[| 6 |
| [50 , 55[| 9 |
| [55, 65[| 9 |

2. La représentation graphique des fréquences sous forme d'un histogramme.

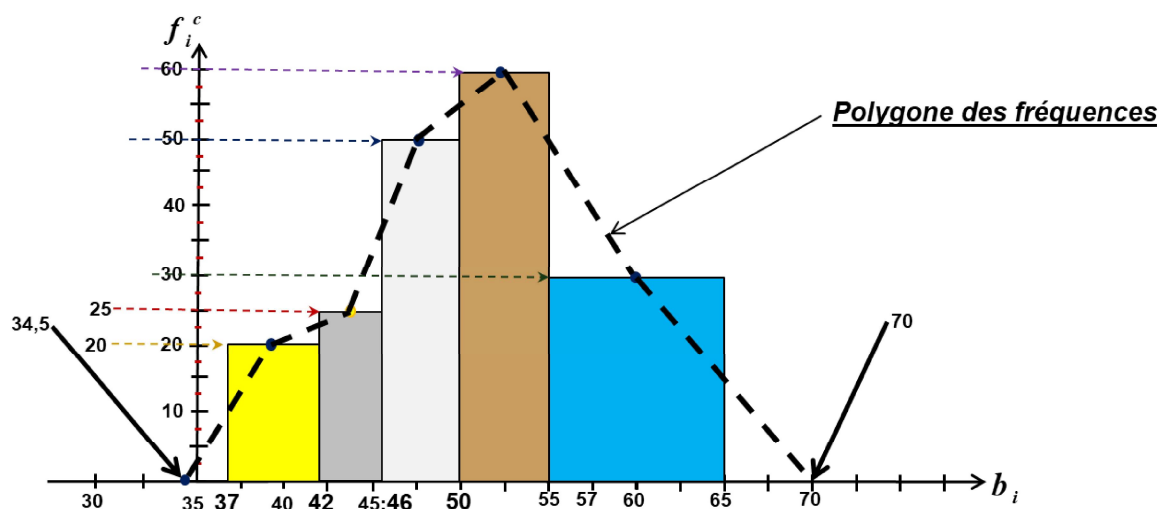
On va calculer les différentes amplitudes :

| Classe | n_i | a_i | $d_i = \frac{n_i}{a_i}$ | $n_i^c = d_i \times a^*$ | f_i | $f_i^c = \frac{n_i^c}{N} = \frac{f_i}{a_i} \times a^*$ |
|--------------|-----------|-------|-------------------------|--------------------------|-------------|--|
| [37, 42[| 3 | 5 | 0,60 | 600 | 0,10 | 20 |
| [42, 46[| 3 | 4 | 0,75 | 750 | 0,10 | 25 |
| [46, 50] | 6 | 4 | 1,50 | 1500 | 0,20 | 50 |
| [50, 55[| 9 | 5 | 1,80 | 1800 | 0,30 | 60 |
| [55, 65[| 9 | 10 | 0,90 | 900 | 0,30 | 30 |
| Total | 30 | | | | 1,00 | |

Donc on obtient l'histogramme des fréquences suivant



Ainsi le polygone des fréquences apparaît sur cet histogramme comme suit



Remarque 1.6.4

1. **Le polygone des effectifs** se définit de la même manière en l'associant à un histogramme des effectifs.
2. Pour le cas d'une variable statistique discrète (**Le polygone des effectifs** respectivement des fréquences) se dessine sur le diagramme en bâton des effectifs (respectivement des fréquences) en reliant les sommets des bâtons.

Exemple 1.6.6 La répartition des notes d'un contrôle noté sur 20 des élèves d'une classe est reportée dans diagramme en bâton des fréquences suivant avec la représentation du polygone des fréquences.

