# Cours de Biostatistique-informatique

***Medouer Nawel***

***Faculté de médecine***

***Université de Batna 2***

*Programme*

***Première partie***

1) *Statistique descriptive univariée*
2) *Statistique descriptive bivariée*
3) *Probabilités*
4) *Variables aléatoires*
5) *Lois de probabilités*

*Programme*

***deuxième partie***

*Statistique inférentielle*

1) *La théorie de l'estimation*
2) *Tests d'hypothèses*
3) *Test de Khi-deux*
4) *Test de l'Anova*

# *Qu'est-ce que la Biostatistique et pourquoi l'étudier?*

## *La Biostatistique*

***La biostatistique*** *ensemble des méthodes qui ont pour objet:*

- ✓ *La collecte des données*
- ✓ *Le traitement des données*
- ✓ *L'interprétation des données*

# *Raisons pour apprendre la biostatistique*

- ✓La statistique s'applique à la plupart des disciplines : médecine, agronomie, biologie, démographie, économie, sociologie, psychologie, . .
- ✓Le but de la statistique est d'extraire des informations pertinentes d'une liste de nombres difficile à l'interpréter par une simple lecture.

## *La statistique descriptive*

***Statistique descriptive*** *branche de la statistique qui a pour but le regroupement des données en utilisant plusieurs techniques*

# *Concepts de base*

*Les observations constituent la source des informations statistiques, avant de débuter l'étude, il faut métriser les notions de base :*

***La population** est l'ensemble que l'on observe et dont chaque élément est appelé **individu** ou **unité statistique.***

***Un échantillon** ou (lot) est une partie (ou sous ensemble) de la population considérée et sa taille correspond à son cardinal.*

## *Concepts de base*

***Le caractère (**ou **variable)** étudié est la propriété observée dans la population ou l'échantillon considéré.*

***Modalités** sont les différentes situations prises par le caractère ou les valeurs possibles de la variable.*

*Les **variables** sont désignées par simplicité par une lettre (X, Y, Z).*

# *Types de variables*

*On distingue deux types de variables :*

***qualitative*** *et* ***quantitative.***

# *Exemple*

***Population :*** *l'ensemble de tous les étudiants inscrits en 1ère année médecine de l'université de Batna.*

***Individu :*** *chaque étudiant.*

***Caractère :*** *l'âge, la taille, la nationalité, la couleur des yeux, …*

***Modalité :*** *17, 18, 19,20 et 21 sont les modalités de la variable l'âge*

*par exemple.*

# *Variables*

- ***Variable qualitative***

  *La variable est dite qualitative quand les modalités ne sont pas mesurables, sont des catégories.*

***On distingue deux types de variables qualitatives***

## *Variable qualitative nominale*

*La variable est dite qualitative nominale quand les modalités ne peuvent pas être ordonnées.*

***Exemple :***

*Groupe sanguin : A, B, O,  AB.*

## *Variable binaire*

*Il s'agit d'un type particulier de variable qualitative nominale qui ne peut prendre que deux modalités.*

*Ce type de variable est extrêmement utilisé dans les sciences de vie et notamment en épidémiologie*.

## *Exemple*

| Variables | modalités |
| --- | --- |
| **sexe** | Homme Femme |
| **Statut vaccinal** | Vacciné non vacciné |
| **Statut dans une étude épidémiologique** | Cas témoin |

# *Variable qualitative ordinale*

- *La variable est dite qualitative ordinale quand les  modalités  peuvent être ordonnées ( s’expriment en modalités qui peuvent être ordonnées selon une échelle*

- ***Exemple : complication d’une maladie***

*(modérée, moyenne, sévère)*

# *Variable quantitative*

*Une variable est dite **quantitative** si toute ses modalités possibles*

*sont des valeurs numériques.*

***On distingue deux types de variables*** quantitatives

## *Variable quantitative discrète*

- Les variables discrètes sont des variables numériques discontinues (des valeurs isolées), Le plus souvent il s'agit de nombres entiers.

- L'ensemble des valeurs possibles est dénombrable.

## *Variable quantitative continue*

- *Une variable est dite **continue**, si l'ensemble des valeurs possibles est continu.*

- ***Exemple:***

  *Poids, Taille, Pression artérielle*

# *Tables statistiques et représentations graphiques*

*Soit une population de taille N (possédant N éléments) sur laquelle on a*

*étudié une variable ayant k valeurs possibles (* $x_1, x_2, \ldots x_k$*)*

*Ces valeurs sont des modalités dans le cas qualitatif*

- $n_i$ : **l'effectif** de la valeur $x_i$ (De la modalité i).
- N : l'effectif total ou taille de la population

$$\boldsymbol{N = \sum_{i=1}^{k} n_i} \, .$$

- On appelle **fréquence relative** de la valeur $x_i$ la

quantité $\boldsymbol{f_i = \frac{n_i}{N}}$

- On a $\boldsymbol{\sum_{i=1}^{k} f_i}$**=1**

- La **proportion** d’individus ayant pris la valeur $x_i$ est noté par
- $\boldsymbol{P_i = \frac{n_i}{N} \times 100 = f_i \times 100}$.

- On a $\sum_{i=1}^{k} \boldsymbol{P_i}$**=100**

# *Variable qualitative*

***Exemple :*** *Parmi les 68 ulcères gastriques colligés :*

- *62 sont de siège antral*
- *2 de siège pré pylorique.*
- *3 de siège sous cardial*
- *1 double ulcère antral et fundique.*

## *Variable qualitative*

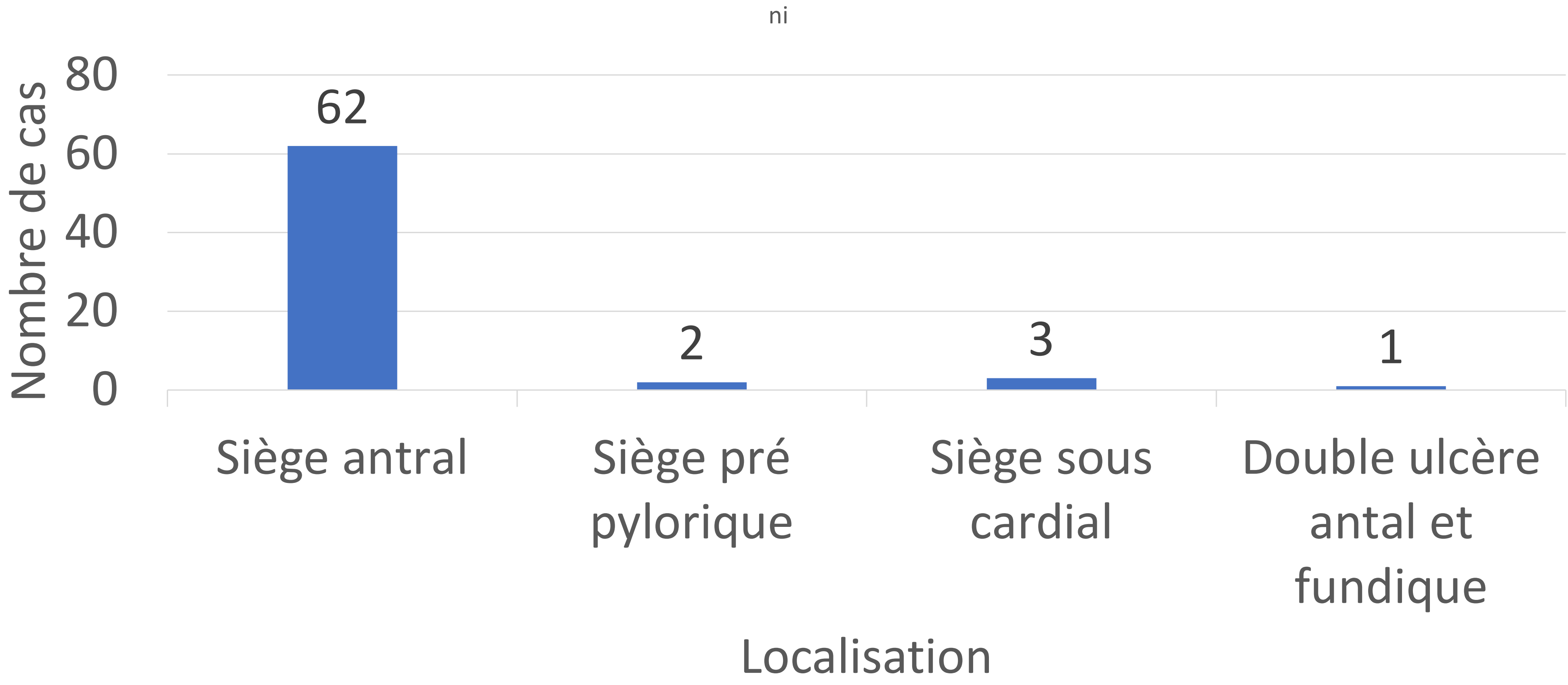
- **Population :** 68 malades (ulcères gastriques)

- **Variable :** Localisation de l’ulcère.

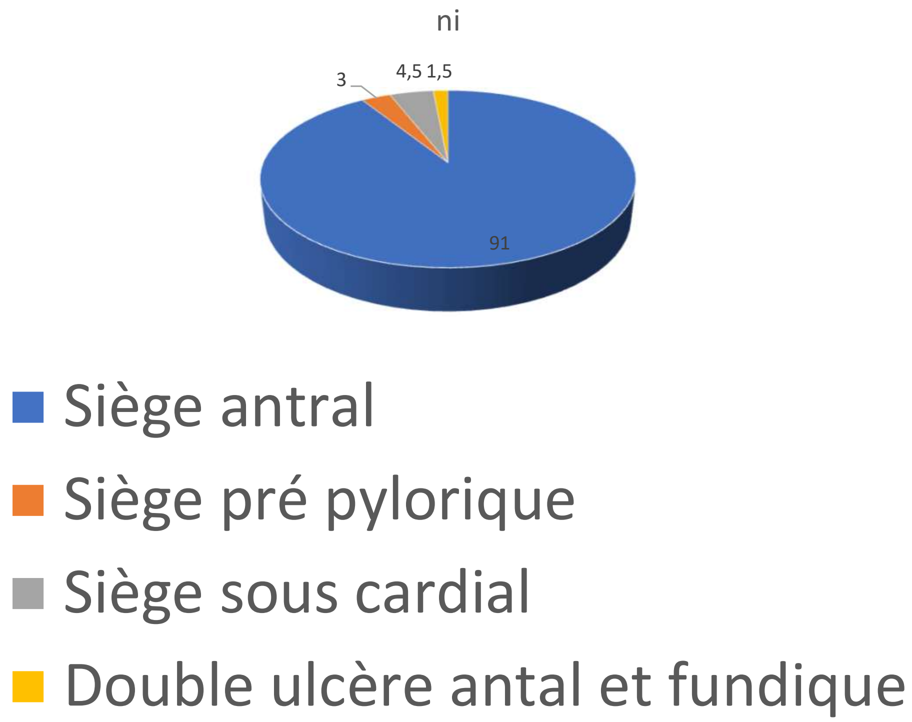
# *Tableau et représentation graphique*

| Localisation<br>Modalités $x_i$ | $n_i$ | $f_i$ | $p_i$ |
| --- | --- | --- | --- |
| **Ulcère antral** | 62 | $\frac{62}{68}$ | 91% |
| **Ulcère pré pylorique** | 2 | $\frac{2}{68}$ | 3% |
| **Ulcère sous cardial** | 3 | $\frac{3}{68}$ | 4.5% |
| **Ulcère double antral et fundique** | 1 | $\frac{1}{68}$ | 1.5% |
| **Total** | 68 | 1 | 100 |

*Figu*re 1.1 – Représentation en tuyaux d’orgue des effectifs (Diagramme en bandes)



# *Figure 1.2 – Représentation en diagramme en secteurs (Camembert)*



# *Diagramme en secteurs des pourcentages*

*Les angles correspondant de l'exemple sont :*

$$\alpha_i = 360° \times f_i$$

# *Variable quantitative*

***Soit des données brute ( sous forme d'une liste)***

***Exemple 2***

*les notes sur 20 de 30 copies d'examen en biostatistique sont :*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *12* | *5* | *11* | *15* | *12* | *19* | *7* | *11* | *7* | *15* |
| *16* | *3* | *3* | *1* | *8* | *7* | *15* | *11* | *7* | *5* |
| *7* | *7* | *19* | *11* | *15* | *1* | *3* | *16* | *3* | *8* |

Soit de données regroupées dans un tableau statistique par une **variable discrète**

# *En 1$^{er}$ lieu, ordonner les données*

**Les données classées par ordre croissant :**

*1 1 3 3 3 3 5 5 7*

*7 7 7 7 8 8 11 11 11 11*

*12 12 15 15 15 15 16 16 19 19*

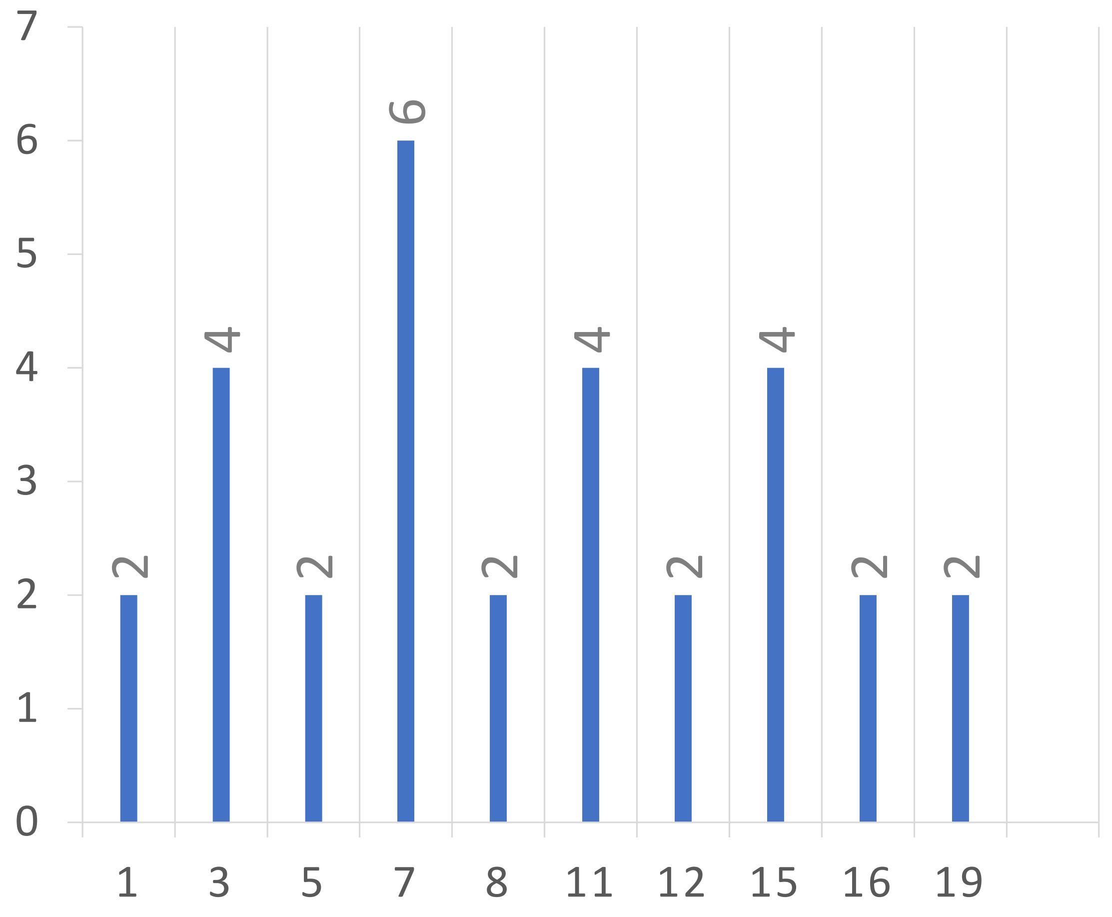
| $x_i$ | $n_i$ |
|---|---|
| 1 | 2 |
| 3 | 4 |
| 5 | 2 |
| 7 | 6 |
| 8 | 2 |
| 11 | 4 |
| 12 | 2 |
| 15 | 4 |
| 16 | 2 |
| 19 | 2 |
| Total | 30 |

## *Représentation graphique*
## *Diagramme en batônnets (en bâtons)*

*Quand la variable est discrète, les effectifs sont présentés par*

*des batônnets (voir Figure1.3)*

# Figure 1.3 – Représentation en diagramme en batônnets (en bâtons)



*Par une* ***variable continue*** *(regroupées en classes)*

***Exemple 4***

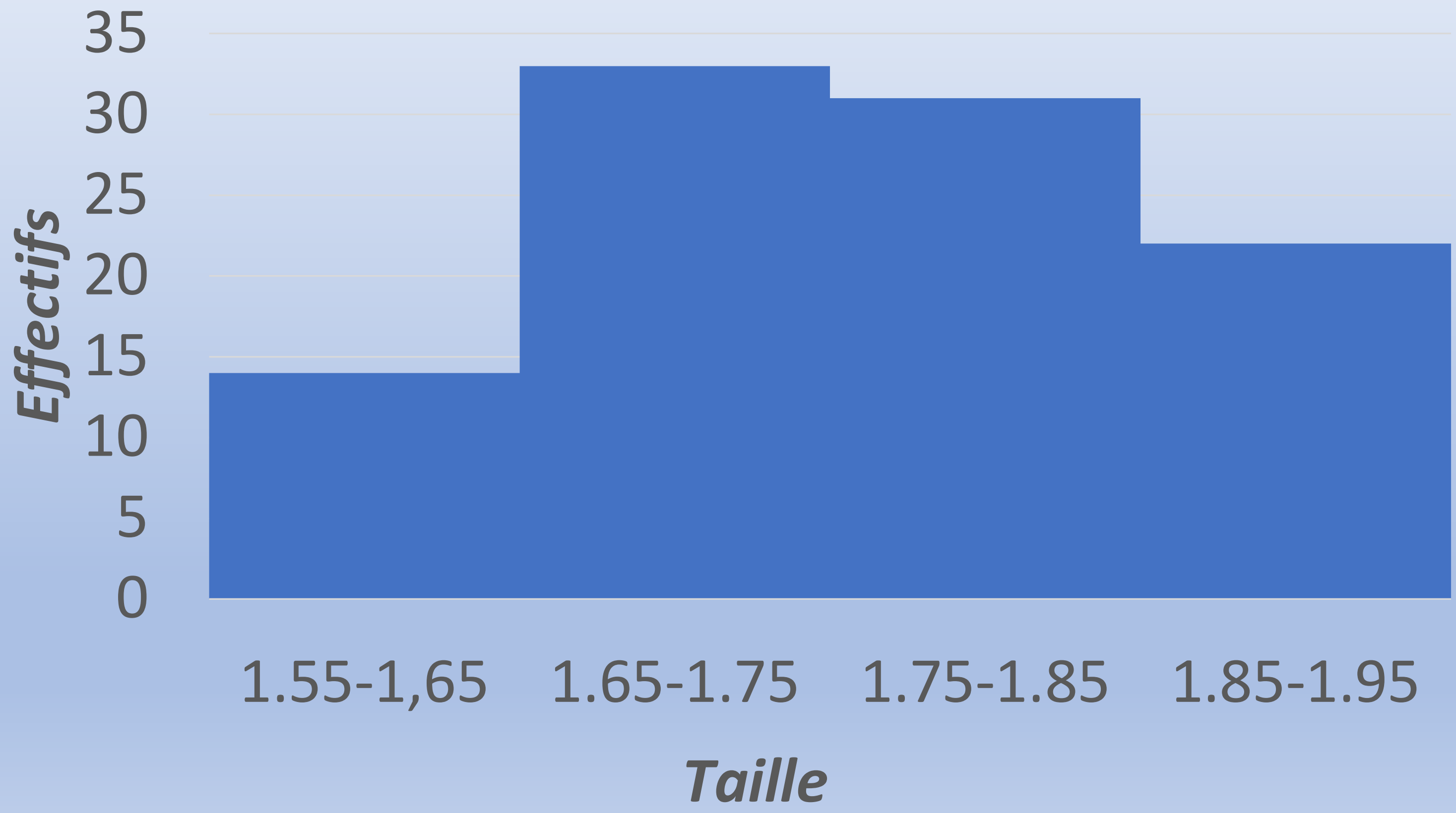
*Une étude faite sur la taille d'un groupe de 100 étudiants*

*(en mètre) a donné les résultats suivants:*

# *Données regroupées en classes*

| Classes $[x_{i-1};\ \ x_i[$ | Effectifs $n_i$ | Centres $c_i$ |
|---|---|---|
| **[1.55 ; 1.65[** | 14 | 1.6 |
| **[1.65 ; 1.75[** | 33 | 1.7 |
| **[1.75 ; 1.85[** | 31 | 1.8 |
| **[1.85 ; 1.95[** | 22 | 1.9 |
| **Total** | 100 | |

# *Figure 1.4 Histogramme*



## *Cas d'une variable statistique continue :*

Lorsque la variable statistique est continue les données sont groupées en classes

$$[x_0, x_1[, [x_1, x_2[,\ldots, [x_{k-1}, x_k[.$$

Les modalités sont sous forme de classes

$$[x_{i-1}, x_i[$$

Ayant les centres $c_i$

# *Variable continue*

$\boldsymbol{x_{i-1}}$ :la borne inférieure de la classe $[x_{i-1}, x_i[$

$x_i$ : la borne supérieure de la classe $[x_{i-1}, x_i[$

$h = x_i - x_{i-1}$ :l'amplitude de la classe $[x_{i-1}, x_i[$

$c_i$ : centre de la classe $[x_{i-1}, x_i[$, avec :

$$c_i = \frac{x_{i-1}+x_i}{2} \qquad \text{pour} \qquad i = \overline{1,k}$$

## ***Calcul des centres de classes***

$$c_i = \frac{x_{i-1} + x_i}{2} = x_{i-1} + {}^{h}/_{2} = x_i - {}^{h}/_{2}$$

*Parce que :*

$$x_{i-1} \underset{h/2}{\longleftrightarrow} c_i \underset{h/2}{\longleftrightarrow} x_i \underset{h/2}{\longleftrightarrow} c_{i+1} \underset{h/2}{\longleftrightarrow} x_{i+1}$$

On obtient

$$c_{i+1} - c_i = h$$

# *Caractéristiques de tendance centrale (Position)*

## *Caractéristiques de tendance centrale*

1. ***La moyenne***
2. ***Le mode***
3. ***La médiane***

## *Caractéristiques de tendance centrale*

1. ***La moyenne*** *: Notée par* $\overline{\boldsymbol{x}}$ *ou encore* ***m*** *est*

*définie comme étant égale à la somme des*

*observations divisée par l'effectif total de la série*

*Données brutes*

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}$$

*Le cas discret*

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i x_i}{N}$$

*Le cas continu*

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i c_i}{N}$$

## *Le mode*

C’est la valeur dominante, la plus fréquente

Ayant l’effectif le plus élevé, Noté par $M_0$

## *Les étapes utilisées pour identifier le* mode *:*

Variable discrète :

- On désigne l'effectif le plus élevé
- La modalité qui correspond à cet effectif c'est exactement le mode.

## *La distribution peut être :*

- *Unimodale (un seul mode)*
- *Bimodale, lorsque deux valeurs distinctes de la variable*

*Statistique correspondent au plus grand effectif à la fois*

- *Multimodale, si la série possède plusieurs modes différents.*

## *Exemples :*

*La série* $\{4, 5, 7{,}7, 11{,}13\ 14\}$ *est unimodale*

*La série* $\{4, 5, 7{,}7{,}7, 11{,}13, 13, 13, 14\}$ *est bimodale, elle a 2 modes différents 7 et 13*

| Modalités | 0 | 2 | 4 | 8 | 10 |
|---|---|---|---|---|---|
| Effectifs | 3 | 5 | 2 | 10→<br>L'effectif le plus élevé | 1 |

**L'effectif le plus élevé c'est 10**

**Donc le mode est de** $M0 = 8$

*Variable continue:*

*Détermination numérique du mode:*

- *On désigne l'effectif le plus élevé*
- *La classe qui correspond à cet effectif c'est exactement la classe modale.*

$$[x_{i-1}, x_i[$$

## Le mode est calculé dans ce cas par cette formule :

$$M_0 = x_{i-1} + h\frac{n_i - n_{i-1}}{2n_i - (n_{i+1} + n_{i-1})}$$

(Admise sans démonstration)

*Tels que:*

- ✓ $x_{i-1}$: la borne inférieure de la classe modale.
- ✓ h : l’amplitude classe
- ✓ $n_{i-1}$ : l’effectif de la classe qui précède la classe modale

- ✓ $n_i$ : l'effectif de la classe modale
- ✓ $n_{i+1}$ : l'effectif de la classe qui suit la classe modale

.

## *Détermination graphique du mode*

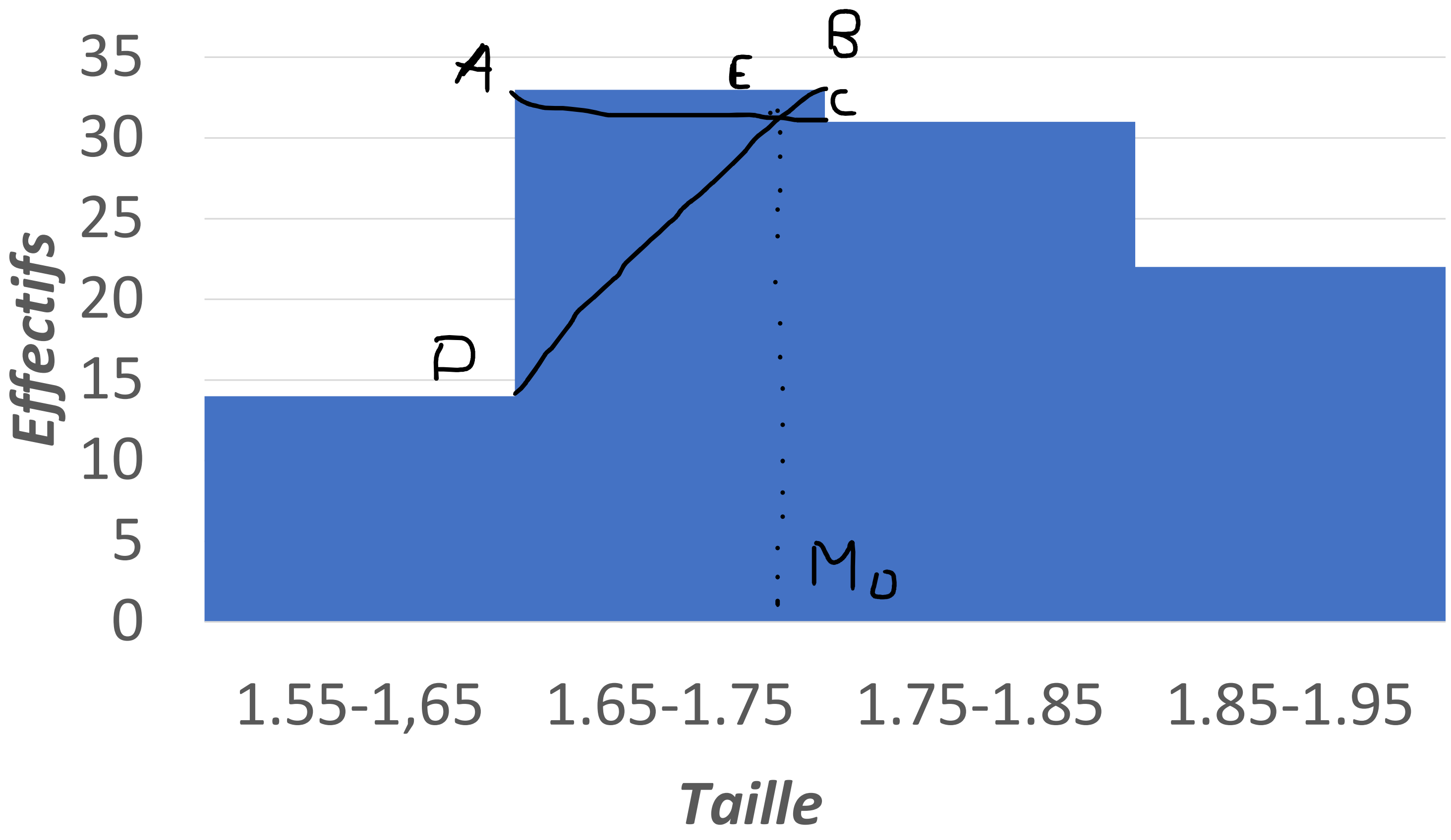
***En utilisant l'histogramme***

***Le mode est l'abscisse du point E d'intersection des diagonales du trapèze ABCD.***

*Exemple*

| Classes $[x_{i-1} ; \; x_i[$ | Effectifs $n_i$ | Centres $c_i$ |
| --- | --- | --- |
| **[1.55 ; 1.65[** | 14 | 1.6 |
| **[1.65 ; 1.75[** | 33 | 1.7 |
| **[1.75 ; 1.85[** | 31 | 1.8 |
| **[1.85 ; 1.95[** | 22 | 1.9 |
| **Total** | 100 | |

# *Figure 1.4 Histogramme*



## *Détermination numérique du mode*

- ✓ *L'effectif le plus élevé c'est 33*
- ✓ *Donc la classe modale :* **[1.65 ; 1.75[**

## *Détermination numérique du mode*

$$M_0 = x_{i-1} + h\frac{n_i - n_{i-1}}{2n_i - (n_{i+1} + n_{i-1})}$$

$$= 1.65 + 0.1\frac{33 - 14}{2 \times 33 - (31 + 14)}$$

$$= 1.74$$

## 3.La médiane

***La médiane, noté par $M_e$ c'est la valeur qui divise la série statistique en deux parties égales, 50% d'observations lui sont inférieures et 50% d'observations lui sont supérieures .***

## 3. La médiane

$$\boldsymbol{x_{min}} \underset{\mathbf{50\%}}{\longleftrightarrow} \boldsymbol{M_e} \underset{\mathbf{50\%}}{\longleftrightarrow} \boldsymbol{x_{max}}$$

# *Détermination numérique de la médiane*

**En 1er lieu, ordonner les données par ordre croissant:**

**Les données classées par ordre croissant :**

*1 2 3 3* $\boxed{3}$ *4 5 5 7*

*Alors la valeur qui divise la série en deux parties égales c'est*

*3, donc la médiane égale à 3*

# *Détermination numérique de la médiane*

***En 1^er^ lieu, il faut d'abord ranger les observations par ordre croissant***

**Les données classées par ordre croissant :**

$\underbrace{1 \quad 2 \quad 3 \quad 3} \quad \boxed{3 \quad 4} \quad \underbrace{5 \quad 5 \quad 7 \quad 9}$

*Alors la valeur qui divise la série en deux parties égales c'est*

$M_e = \frac{3+4}{2} = 3{,}5$ *donc la médiane égale à 3,5*

## Pourquoi la détermination défère de l’exemple1 à l’exemple 2 ?

***Parce que la médiane dépend de l’effectif totale N***

***c’est-à-dire une série paire ou impaire***

***Alors pour une série impaire vous avez une seule valeur centrale***

***par contre pour une série paire vous avez deux***

***Valeurs centrales***

## *Cas général*

$N\ impair$

$$M_e = (\frac{N+1}{2})^{\text{ième}} \text{ observation}$$

$N\ pair$

$$M_e = \frac{(\frac{N}{2})^{\text{ième}}\ obsevation + (\frac{N}{2}+1)^{\text{ième}} obsevation}{2}$$

## *Remarques*

- ***La médiane se trouve au milieu de la série (centre)***
- ***La médiane n’est pas liée à la valeur numérique des observations mais à leur position***
- ***La médiane n’est pas toujours parmi les observations***

## *Série groupée dans un tableau statistique*
## *Variable discrète*

***On débute toujours par le calcul des effectifs cumulés $n_i\uparrow$ croissants***

***Comment?***

$n_1\uparrow = n_1$

$n_2\uparrow = n_1 + n_2$

$n_3\uparrow = n_1 + n_2 + n_3$

.

.

.

$n_k\uparrow = n_1 + n_2 + \cdots + n_k$

*Série groupée dans un tableau statistique*
*Variable discrète*

Si la série est paire la médiane

**N impair**

$$M_e = \left(\frac{N+1}{2}\right)^{ième} \text{ observation}$$

*Série groupée dans un tableau statistique*

*Variable discrète*

Si la série est impaire

**N pair**

$$M_e = \frac{(\frac{N}{2})^{\text{ième}}\ observation + (\frac{N}{2}+1)^{\text{ième}} observation}{2}$$

| Modalités $x_i$ | Effectifs $n_i$ | Effectifs cumulés $n_i \uparrow$ |
|---|---|---|
| ***0*** | 4 | 4 |
| 1 | 14 | 18 |
| 2 | 10 | 28 |
| 3 | 15 | 43 |
| 4 | 7 | 50 |
| total | 50 | |

## *Exemple*

**N = 50   pair**

$$M_e = \frac{(\frac{N}{2})^{ième}\ observation + (\frac{N}{2}+1)^{ième} observation}{2}$$

$$M_e = \frac{(\frac{50}{2})^{ième}\ observation + (\frac{50}{2}+1)^{ième} observation}{2}$$

| Modalités $x_i$ | Effectifs $n_i$ | Effectifs cumulés $n_i \uparrow$ |
|---|---|---|
| ***0*** | 4 | 4 |
| 1 | 14 | 18 |
| 2 | 10 | $\rightarrow (25)^{ième}\ observa$<br>$\rightarrow (26)^{ième}\ observa$<br>28 |
| 3 | 15 | 43 |
| 4 | 7 | 50 |
| total | 50 | |

*Exemple*

Si la série est paire la médiane

$$M_e = \frac{(25)^{ième}\ observation + (26)^{ième} observation}{2}$$

$$= \frac{2+2}{2} = 2$$

| Modalités $x_i$ | Effectifs $n_i$ | Effectifs cumulés $n_i \uparrow$ |
|---|---|---|
| ***0*** | 4 | 4 |
| 1 | 14 | 18 |
| 2 | 10 | $\rightarrow$ $(26)^{ième}$ $observati$ 28 |
| 3 | 16 | 44 |
| 4 | 7 | 51 |
| total | 51 | |

*Exemple*

*Si la série est impaire la médiane*

**N impair**

$$M_e = \left(\frac{N+1}{2}\right)^{\text{ième}} \text{ observation}$$

$$M_e = (26)^{\text{ième}} observation$$

=2

*Variable continue:*

*Détermination numérique du médiane:*

- *On calcule les effectifs cumulés.*
- *On désigne d'abord la classe médiane*

$$[x_{i-1}, x_i[$$

*Comment?*

*Variable continue:*

*Détermination numérique du médiane:*

- *La classe médiane est la classe contenant l'observation d'ordre* $\frac{N}{2}$ *si La série est paire.*
- *La classe médiane est la classe contenant l'observation d'ordre* $\frac{N+1}{2}$ *si La série est impaire.*

**La médiane est calculée dans ce cas par cette formule :**

$$M_{\mathrm{e}} = x_{i-1} + h\,\frac{\frac{N}{2} - n_{i-1}\uparrow}{n_i\uparrow - n_{i-1}\uparrow}$$

(Admise sans démonstration)

*Tels que:*

- ✓$x_{i-1}$: la borne inférieure de la classe médiane.
- ✓h : l'amplitude classe
- ✓$n_i \uparrow$ : l'effectif cumulé de la classe médiane

✓ $n_{i-1} \uparrow$ : l'effectif cumulé de la classe qui précède la classe médiane,

## *Détermination graphique du médiane*

***En utilisant la courbe cumulative croissante, la médiane est l'abscisse du point*** $\boldsymbol{\frac{N}{2}}$**.**

## *La courbe cumulative croissante F*

$F(x_0) = 0$

$F(x_1) = n_1 \uparrow$

$F(x_2) = n_2 \uparrow$

$F(x_3) = n_3 \uparrow$

.

.

.

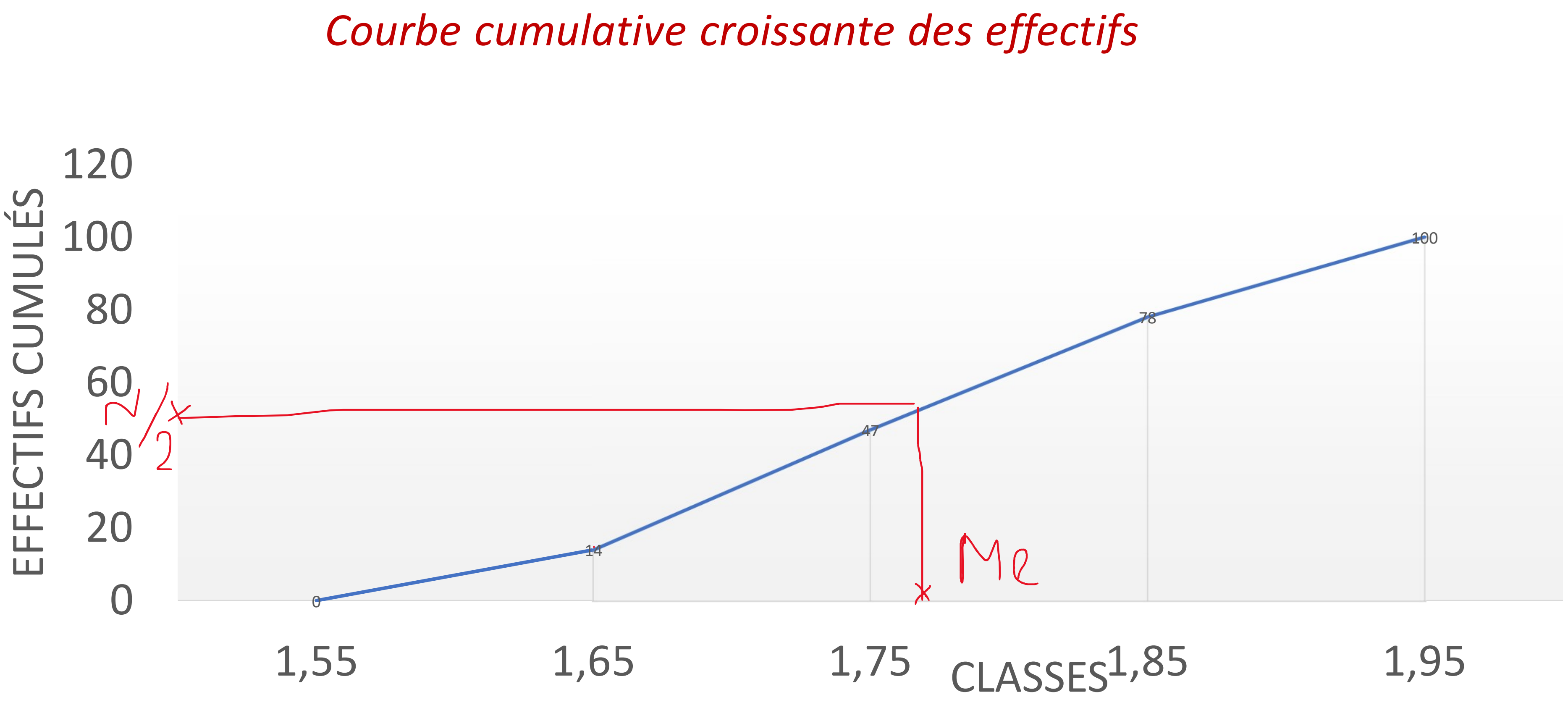
$F(x_k) = n_k \uparrow$

*La courbe cumulative croissante F*

$On\ joint\ les\ points\ par\ des\ segments$

La médiane est exactement l'abscisse du point $\frac{N}{2}$

# Détermination graphique du médiane



# *Caractéristiques de Dispersion (Indicateurs)*

# Caractéristiques de dispersion

1. ***L'étendue***
2. ***L'intervalle interquartile***
3. ***La variance***
4. ***L'écart-absolu***
5. ***Le coefficient de dispersion***

# *L'étendue*

***Variable discrète***

$$e = max\,(x_i) - min(x_i)$$

***Variable continue***

$$e = x_k - x_0$$

## La variance

## Le cas discret

$$Var = \frac{\sum_{i=1}^{K} n_i x_i^2}{N} - m^2 = \frac{\sum_{i=1}^{K} n_i (x_i \ -m)^2}{N}$$

## Le cas continu

$$Var = \frac{\sum_{i=1}^{k} n_i c_i^2}{N} - m^2 = \frac{\sum_{i=1}^{K} n_i (c_i \ -m)^2}{N}$$

# L'écart-type

$$\sigma = \sqrt{Var}$$

## Le cas discret

$$\sigma = \sqrt{\frac{\sum_{i=1}^{K} n_i x_i^2}{N} - m^2}$$

### Le cas continu

$$\sigma = \sqrt{\frac{\sum_{i=1}^{k} n_i c_i^2}{N} - m^2}$$

## *Coefficient de variation* $\boldsymbol{Cv}$

$$\mathrm{Cv} = \frac{\sigma}{\mathrm{m}}(\times 100)$$

Le coefficient de variation est un indicateur

de dispersion

## Interprétation de coefficient de variation $\boldsymbol{Cv}$

$Cv < 0{,}33$ *(33%)*

*Valeurs concentrées autour de la moyenne*

$Cv$*>0,33 (33%)*

*Série dispersée autour de la moyenne*

## *Caractéristiques de dispersion*

*On utilise souvent les paramètres de dispersion pour comparer la dispersion de deux distributions statistiques*

*La distribution ayant un paramètre plus élevé est la* plus dispersée.

## *Coefficient de variation $\boldsymbol{Cv}$*

*Remarque on utilise souvent le coefficient de variation*

*pour  comparer deux distributions ayant deux unités de*

*mesure différentes, Parce que ce dernier s'exprime sans unité de mesure.*

*La distribution ayant un coefficient plus élevé c'est la* plus dispersée.

## *Les quartiles*

On a définit la médiane pour répartir la population en moitié. Pour la répartir en quarts, on définit des paramètres appelés les quartiles.

Ils sont au nombre de trois et sont notés par $Q_1, Q_2, Q_3$ .

Ce sont donc des valeurs partageants, en quatre parties d'effectifs égaux

# *Intervalle interquartile*

$$I_Q = Q_3 - Q_1$$

Le ***1^er^ quartile***

$$\boldsymbol{x_{min} \underset{25\%}{\longleftrightarrow} Q_1 \underset{75\%}{\longleftrightarrow} x_{max}}$$

*Le 3ème quartile*

$$x_{min} \underset{75\%}{\longleftrightarrow} Q_3 \underset{25\%}{\longleftrightarrow} x_{max}$$

Le deuxième quartile n'est d'autre que la médiane

$$Q_2 = M_e$$

# *Comment déterminer les quartiles?*

On procède la même procédure que pour la médiane

***Cas discret:***

✓ ***On calcule*** $\boldsymbol{\frac{N}{4}}$ *pour le 1er quartile*

*(respectivement* $\frac{3N}{4}$ *pour le 3ème quartile)*

## *Comment déterminer les quartiles?*

On procède la même procédure que pour la médiane

***Cas discret:***

✓ ***Si*** $\boldsymbol{\frac{N}{4}}$ *est un entier (respectivement* $\frac{3N}{4}$ *est un entier)*

*Alors le 1er quartile(respectivement le 3ème quartile)*

*est l'observation d'ordre* $\boldsymbol{\frac{N}{4}}$ *(respectivement* $\frac{3N}{4}$*)*

## *Comment déterminer les quartiles?*

On procède la même procédure que pour la médiane

***Cas discret:***

✓ ***Si*** $\boldsymbol{\frac{N}{4}}$ *n'est pas un entier (respectivement* $\frac{3N}{4}$ *n'est pas un entier)*

*Alors le 1er quartile(respectivement le 3ème quartile)*

*Est l'observation d'ordre* $\boldsymbol{\grave{N}}$

*Tel que* $\boldsymbol{\grave{N}}$ *est l'entier qui suit* $\boldsymbol{\frac{N}{4}}$ *(respectivement* $\frac{3N}{4}$*)*

# *Exemple:*

| Modalités | 8 | 12 | 16 | 20 | 24 | 28 | Total | |
|---|---|---|---|---|---|---|---|---|
| Effectifs | 7 | 20 | 23 | 19 | 14 | 6 | | |
| Effectifs cumulés | 7 | 27 | 50 | 69 | 83 | 89 | | |

# *Détermination des quartiles*

- $\frac{N}{4} = 22{,}25$ *(n'est pas un entier, on passe directement à l'entier suivant)*
- ***Le 1^er^ quartile correspond à l'observation d'ordre 23***

- $\frac{3N}{4} = 66{,}75$ *(n'est pas un entier, on passe directement à l'entier suivant)*

- ***Le 3ème quartile correspond à l'observation d'ordre 67***

## ***Les quartiles sont calculés dans le cas continu par les formules :***

$$Q_1 = x_{i-1} + h\frac{\frac{N}{4} - n_{i-1}\uparrow}{n_i\uparrow - n_{i-1}\uparrow}$$

$$Q_3 = x_{i-1} + h\frac{\frac{3N}{4} - n_{i-1}\uparrow}{n_i\uparrow - n_{i-1}\uparrow}$$

# *Détermination graphique des quartiles*

***En utilisant la courbe cumulative croissante***

***Le 1er quartile est l'abscisse du point $\frac{N}{4}$.***

***Le 3ème quartile est l'abscisse du point***

$\frac{3N}{4}$.

*Exemple*

| Classes $[x_{i-1}\ ;\ x_i[$ | Effectifs $n_i$ | Centres | Effectifs cumulés |
|---|---|---|---|
| [1.55 ; 1.65[ | 14 | 1.6 | 14 |
| [1.65 ; 1.75[ | 33 | 1.7 | 47 |
| [1.75 ; 1.85[ | 31 | 1.8 | 78 |
| [1.85 ; 1.95[ | 22 | 1.9 | 100 |
| Total | 100 | | |

# *Exemple*

$$\frac{N}{4} = 25 \qquad Q_1 \in [1{,}65; 1{,}75[$$

$$Q_1 = x_{i-1} + h\frac{\frac{N}{4} - n_{i-1}\uparrow}{n_i\uparrow - n_{i-1}\uparrow}$$

$$= 1{,}65 + 0{,}1\frac{25-14}{47-14}$$

$$= 1{,}683$$

## *Exemple*

$\frac{N}{2} = 50 \;\; M_e \in [1{,}75; 1{,}85[$

$$M_{\mathrm{e}} = x_{i-1} + h\frac{\frac{N}{2} - n_{i-1}\uparrow}{n_i\uparrow - n_{i-1}\uparrow}$$

$$=1{,}75 + 0{,}1\frac{50-47}{78-47}$$

$$=1{,}7596$$

## *Exemple*

$$\frac{3N}{4} = 75 \quad Q_3 \in [1{,}75; 1{,}85[$$

$$Q_3 = x_{i-1} + h\,\frac{\frac{3N}{4} - n_{i-1}\uparrow}{n_i\uparrow - n_{i-1}\uparrow}$$

$$= 1{,}75 + 0{,}1\frac{75-47}{78-47}$$

$$= 1{,}84$$

# *Détermination graphique du médiane*