

# RAPPEL

## HISTOGRAMME

## EXERCICE: REPRÉSENTEZ GRAPHIQUEMENT L'HISTOGRAMME DES FRÉQUENCES RELATIVES CUMULÉES

Classe	Centre de classe	$n_j$	$f_j$ ( en %)	$N_j$	$F_j$
[0-20000[	10000	20	0.444	20	0.444
[20000-40000[	30000	9	0.200	29	0.644
[40000-60000[	50000	11	0.244	40	0.888
[60000-80000[	70000	3	0.067	43	0.955
[80000-100000[	90000	1	0.022	44	0.977
[100000-120000[	110000	1	0.022	45	1

# GRAPHIQUE SUR BASE DES QUANTILES

- Étape 1 = Déterminer les valeurs des quantiles
- Étape 2 = Représentation graphique (boîtes à moustaches)

# TYPE DE VARIABLES QUE L'ON PEUT REPRÉSENTER VIA DES BOXPLOT

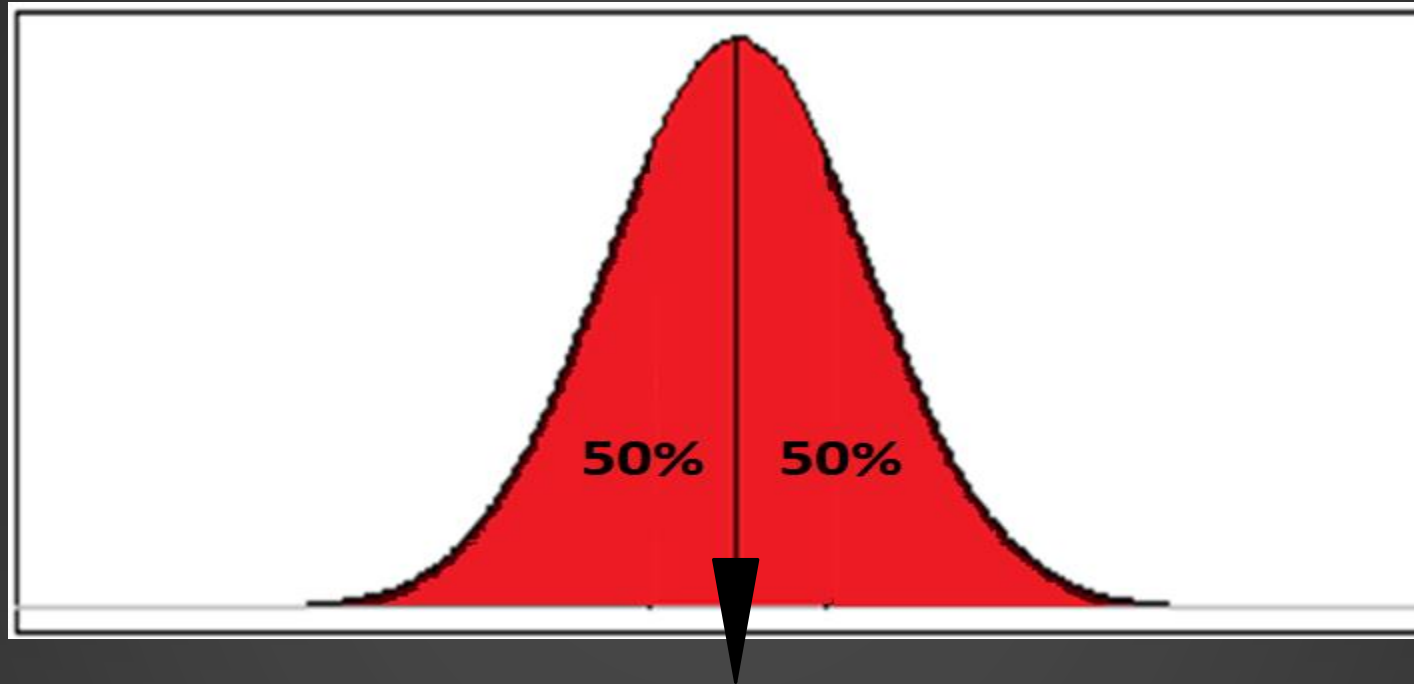
- Variables quantitatives discrètes
- Variables quantitatives continues

# QUANTILES

- Division de la distribution en un certain nombre d'intervalles égaux
  - **La médiane ( $\theta$ )**: division de l'aire sous la courbe en deux portions égales (50%)
  - **Les quartiles**: division de la distribution en 4 portions égales (25%)
  - **Les déciles**: division de la distribution en 10 portions égales (10%)
  - **Les percentiles**: division de la distribution en 100 parties égales (1%)

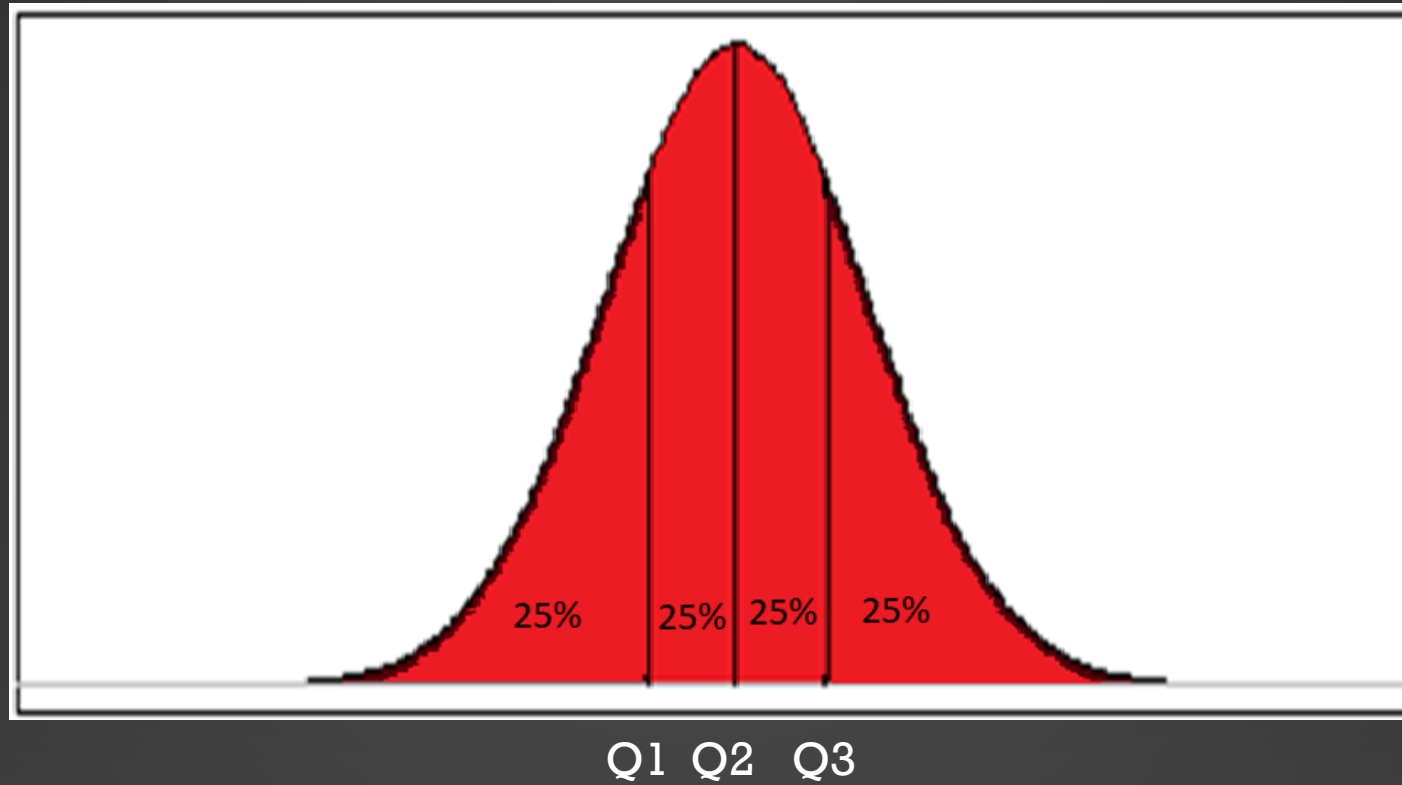


# MÉDIANE



Moyenne  
= médiane  
= 2<sup>ème</sup> quartile  
= 5<sup>ème</sup> décile  
= 50<sup>ème</sup> percentile

# LES QUARTILES





# CALCUL DE LA MÉDIANE ET DES QUARTILES

- Le calcul sera différent suivant qu'on veuille déterminer la médiane sur base:
  - D'une série statistique
  - OU d'une distribution de fréquence
    - Variable discrète (*ex.: le nombre d'enfants*)
    - Variable continue (*ex.: la taille, le poids...*)

# CALCUL EN UTILISANT UNE SÉRIE STATISTIQUE

- Exemple: 3, 7, 2, 11, 9, 8, 1, 13, 15
- *Puisque la médiane divise la distribution en deux parts égales, on voudra autant de données à gauche (= les données les plus basses) et à droite (les données les plus hautes) de la médiane.*

# DÉTERMINER LA MÉDIANE EN 4 ÉTAPES

- 3, 7, 2, 11, 9, 8, 1, 13, 15
- 1, 2, 3, 7, 8, 9, 11, 13, 15
- $n = 9 \rightarrow \text{rang médian} = 5$
- Médiane = 8

# CALCUL EN UTILISANT UNE SÉRIE STATISTIQUE

- Exemple: 1, 2, 3, 7, 8, 9, 11, 13, 15
- Médiane

Premier quartile	Troisième quartile
1, 2, 3, 7	9, 11, 13, 15
Rang quartile = $5/2 = 2.5$	Rang quartile = $5/2 = 2.5$
→ Quartile = $(2+3)/2 = 2.5$	→ Quartile = $(11+13)/2 = 12$

# CALCUL EN UTILISANT UNE SÉRIE STATISTIQUE

- Exemple: 3, 7, 2, 6, 9, 8, 11, 14, 13, 15
- *Puisque la médiane divise la distribution en deux parts égales, on voudra autant de données à gauche (= les données les plus basses) et à droite (les données les plus hautes) de la médiane.*

# DÉTERMINER LA MÉDIANE EN 4 ÉTAPES

- 3, 7, 2, 6, 9, 8, 11, 14, 13, 15
- 2, 3, 6, 7, 8, 9, 11, 13, 14, 15
- $n = 10 \rightarrow \text{rang médian} = 5.5$
- Médiane = 8.5



# CALCUL EN UTILISANT UNE SÉRIE STATISTIQUE

Médiane

- Exemple: 2, 3, 6, 7, 8 | 9, 11, 13, 14, 15

Premier quartile	Troisième quartile
2, 3, 6, 7, 8	9, 11, 13, 14, 15
Rang quartile = $6/2 = 3$	Rang quartile = $6/2 = 3$
→ Quartile = 6	→ Quartile = 13

# DÉTERMINER LA MÉDIANE EN 4 ÉTAPES

- **Étape 1:** ordonner la série statistique
- **Étape 2:** attribuer un rang à chaque observation
- **Étape 3:** déterminer le rang médian:  $(n+1)/2$
- **Étape 4:** déterminer la valeur qui occupe le rang médian

## DÉTERMINER LE PREMIER (OU TROISIÈME) QUARTILE

- **Étape 1**: ne sélectionner QUE les données à gauche (ou à droite) de la médiane
- **Reproduire les mêmes étapes que vues précédemment, pour le calcul de la médiane, sur cette sous-partie des données**

## CALCUL EN UTILISANT UN TABLEAU DE FRÉQUENCES (VARIABLES DISCRÈTES)

$x_j$	$n_j$	$f_j$	$N_j$	$F_j$
1	1	0.11	1	0.11
2	1	0.11	2	0.22
3	1	0.11	3	0.33
7	1	0.11	4	0.44
8	1	0.11	5	0.56
9	1	0.11	6	0.67
11	1	0.11	7	0.78
13	1	0.11	8	0.89
15	1	0.11	9	1

## CALCUL EN UTILISANT UN TABLEAU DE FRÉQUENCES (VARIABLES DISCRÈTES)

$x_j$	$n_j$	$f_j$	$N_j$	$F_j$
1	1	0.11	1	0.11
2	1	0.11	2	0.22
3	1	0.11	3	0.33
7	1	0.11	4	0.44
8	1	0.11	5	0.56
9	1	0.11	6	0.67
11	1	0.11	7	0.78
13	1	0.11	8	0.89
15	1	0.11	9	1

## Commentaire

Par convention, si la proportion correspondant au quantile est exactement atteinte pour une valeur de la variable, le quantile est défini par convention comme la moyenne entre cette valeur et la valeur suivante.



## CALCUL EN UTILISANT UN TABLEAU DE FRÉQUENCES (VARIABLES DISCRÈTES)

$x_j$	$n_j$	$f_j$	$N_j$	$F_j$
2	1	0.10	1	0.10
3	1	0.10	2	0.20
6	1	0.10	3	0.30
7	1	0.10	4	0.40
8	1	0.10	5	0.50
9	1	0.10	6	0.60
11	1	0.10	7	0.70
13	1	0.10	8	0.80
14	1	0.10	9	0.90
15	1	0.10	10	1.00

## CALCUL EN UTILISANT UN TABLEAU DE FRÉQUENCES (VARIABLES DISCRÈTES)

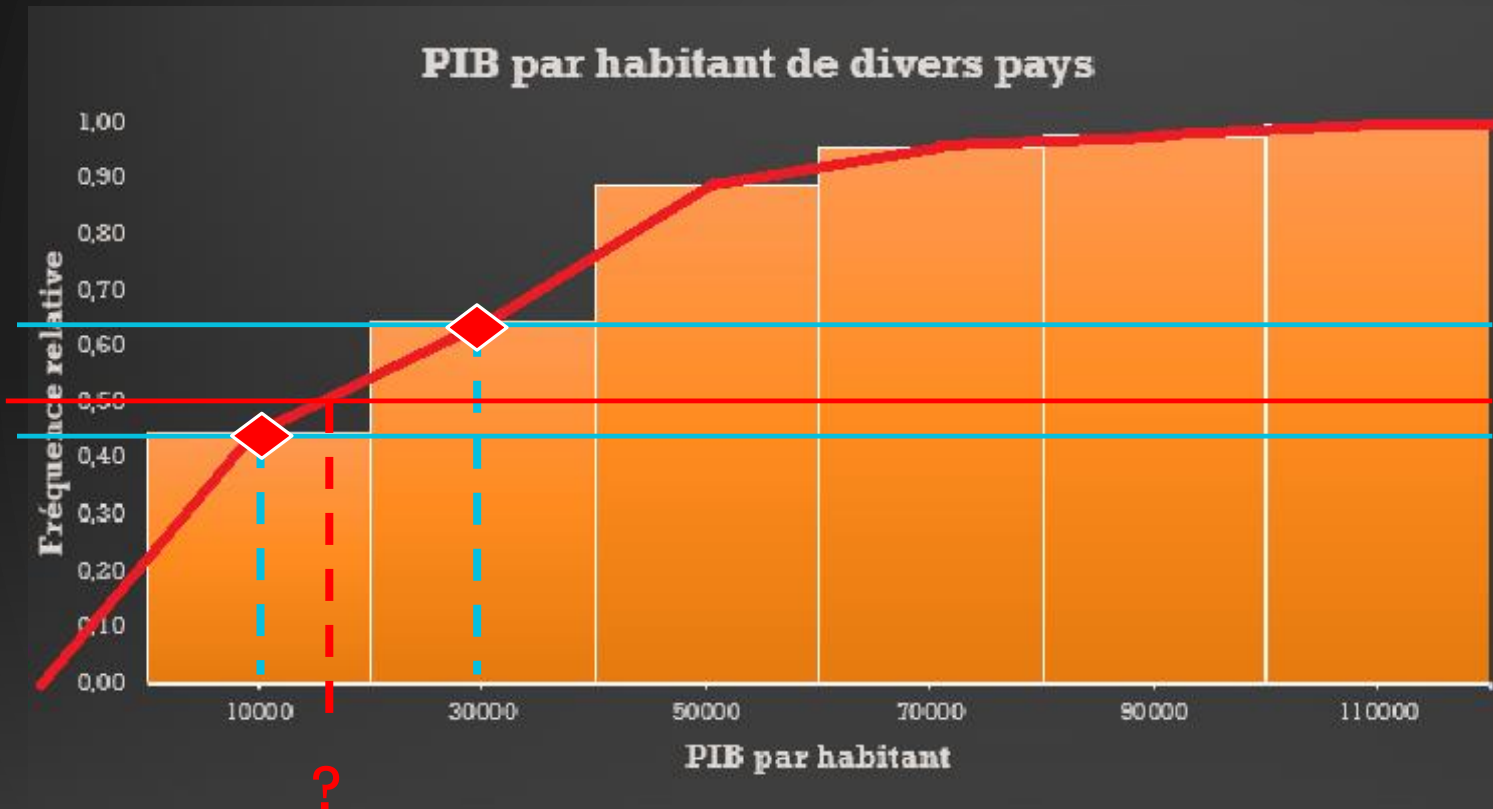
$x_j$	$n_j$	$f_j$	$N_j$	$F_j$
2	1	0.10	1	0.10
3	1	0.10	2	0.20
6	1	0.10	3	0.30
7	1	0.10	4	0.40
8	1	0.10	5	0.50
9	1	0.10	6	0.60
11	1	0.10	7	0.70
13	1	0.10	8	0.80
14	1	0.10	9	0.90
15	1	0.10	10	1.00

- Si variable continue:

Alors, on pourra trouver la valeur exacte des quartiles,  
par **interpolation linéaire**.

# MÉTHODE D'INTERPOLATION LINÉAIRE

- Grâce au polygone de fréquences cumulées.



$$10000 + 5600 = 15600$$

Classe	Centre de classe	$n_j$	$f_j$ (en %)	$N_j$	$F_j$
[0-20000[	10000	20	0.444	20	0.444
[20000-40000[	30000	9	0.200	29	0.644
[40000-60000[	50000	11	0.244	40	0.888
[60000-80000[	70000	3	0.067	43	0.955
[80000-100000[	90000	1	0.022	44	0.977
[100000-120000[	110000	1	0.022	45	1

$$\begin{aligned} &.200 = 20000 \\ &.01 = 1000 \\ &.056 = 5600 \end{aligned}$$

## ERRATUM COURS, P. 27

- La fréquence cumulée associée à la valeur 2 vaut 0.22. La fréquence cumulée associée à la valeur 3 vaut 0.33. ~~La médiane~~ **Le premier quartile** vaudra donc 3.
  - La fréquence cumulée associée à la valeur 9 vaut 0.67. La fréquence cumulée associée à la valeur 11 vaut 0.78. ~~La médiane~~ **Le troisième quartile** vaudra donc 11.
- 
- La fréquence cumulée associée à la valeur 3 vaut 0.2. La fréquence cumulée associée à la valeur 6 vaut 0.3. ~~La médiane~~ **Le premier quartile** vaudra donc 6.
  - La fréquence cumulée associée à la valeur 11 vaut 0.7. La fréquence cumulée associée à la valeur 13 vaut 0.8. ~~La médiane~~ **Le troisième quartile** vaudra donc 13.