

Statistiques

I Série statistique

Définition :

Dans une série de données, l'**effectif** d'une donnée représente le nombre de fois que cette donnée apparaît.

Définition :

L'**effectif total**, noté N , est le nombre d'individus dans la population observée. (c'est la somme des effectifs)

Définition :

La **fréquence** d'une valeur est le quotient de l'effectif de cette valeur par l'effectif total, $f_i = \frac{n_i}{N}$.

Remarques :

- Une fréquence est toujours comprise entre 0 et 1 : $0 \leq f_i \leq 1$.
- En théorie, la somme des fréquence est égale à 1 : $f_1 + \dots + f_p = 1$, mais en pratique, elle n'est souvent que « proche » de 1 (à cause des arrondis).

Définition :

Lorsque les données sont en grand nombre, on peut les regrouper dans des intervalles de la forme $[a ; b[$, appelés **classes**.

$b - a$ est l'**amplitude** de la classe et $\frac{a+b}{2}$ est le **centre** de la classe.

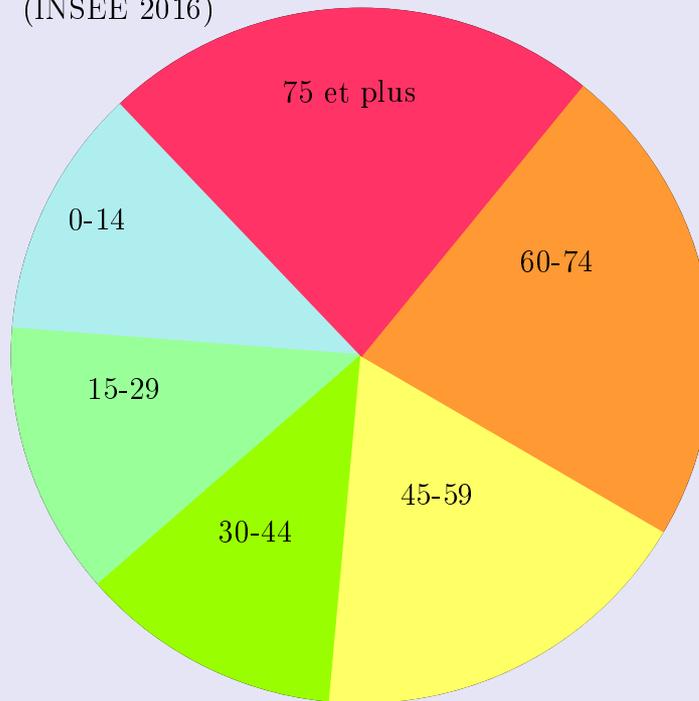
II Représentations des données

Il existe différentes façons d'illustrer des données, plus ou moins pratiques en fonction de ce que l'on veut observer. On peut fabriquer un tableau de données, si on veut pouvoir les lire plus simplement on peut construire des graphiques.

Voici quelques exemples :

II.1 Diagrammes à secteurs

Population par grandes tranches d'âges à La Châtre (INSEE 2016)



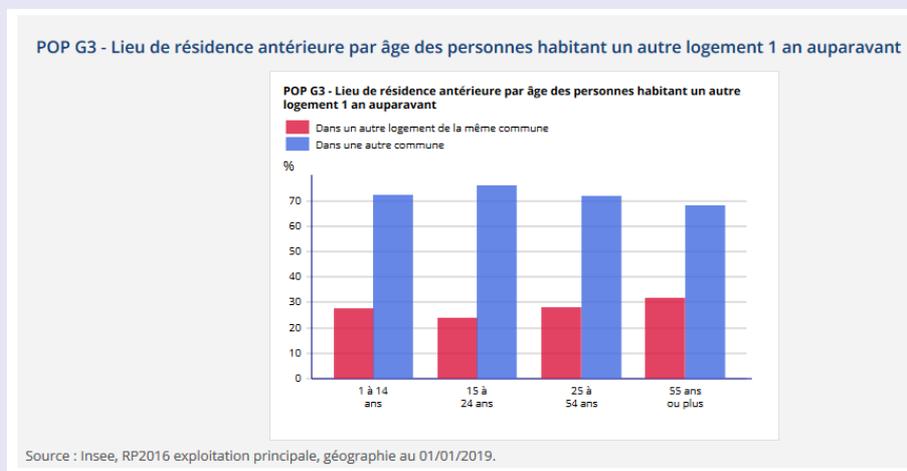
Remarques :

- On dit aussi diagramme circulaire ou camembert.
- L'angle du secteur angulaire est proportionnel à l'effectif de la donnée.
- Il existe aussi des diagrammes semi-circulaires.



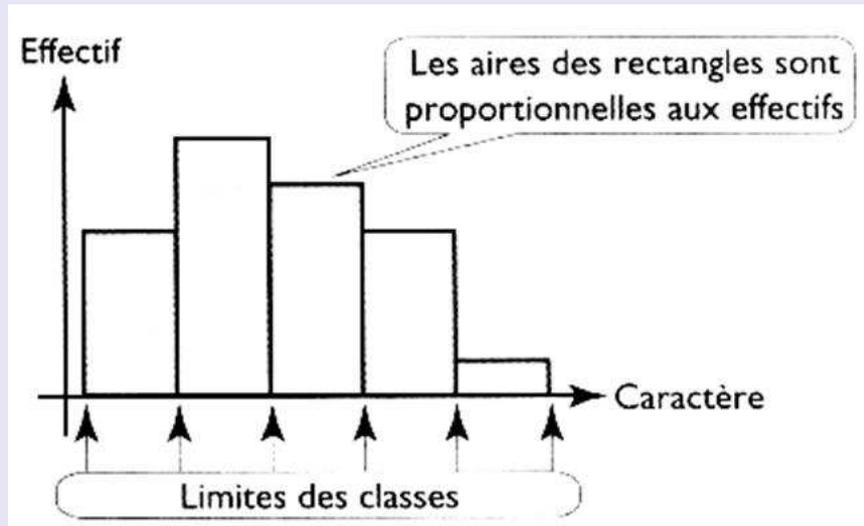
exo 1

II.2 Diagramme en bâtons



exo 2

II.3 Histogramme



III Caractéristiques de position

III.1 Moyenne

Définition :

La moyenne \bar{x} d'une série statistique se calcule :

$$\bar{x} = \frac{n_1 \times x_1 + \dots + n_p \times x_p}{N}$$

où les x_i sont les données, les n_i sont les effectifs et N est l'effectif total.

Exemple :

La **moyenne** de la série 4 ; 5 ; 5 ; 6 ; 6 ; 6 ; 7 ; 7 se calcule :

$$\bar{x} = \frac{1 \times 4 + 2 \times 5 + 3 \times 6 + 2 \times 7}{8}$$

Remarque :

Lorsque les données sont regroupées par classes, on considère que les individus d'une même classe seront représentés par le centre de celle-ci. Avec cette approximation, on se ramène à calculer une moyenne pondérée.

 Exercice : calcul de moyenne

III.2 Médiane

Définition :

Une **médiane** d'une série statistique est une valeur vérifiant que :

- au moins la moitié des valeurs lui soient inférieures ou égales ;
- au moins la moitié des valeurs lui soient supérieures ou égales.

Exemples :

Considérons les notes (classées dans l'ordre croissant) de deux élèves.

- Albert : 8 ; 9 ; 11 ; 12 ; 16.

La note médiane est 11.

- Bertrand : 10 ; 10 ; 11 ; 12 ; 13 ; 16.

La note médiane est la moyenne des deux valeurs centrales : $\frac{11+12}{2} = 11,5$.

 Exercice : calcul de médiane

IV Caractéristique de dispersions

IV.1 Étendue

Définition :

L'**étendue** d'une série statistique est la différence entre la plus grande valeur et de la plus petite valeur de la série.

Exemples :

En reprenant les notes d'Albert et de Bertrand ci-dessus :

- Albert : L'étendue de sa série de note est 8 : $16 - 8 = 8$.

- Bertrand : L'étendue de sa série de note est 6 : $16 - 10 = 6$.

 Vidéo : calcul moyenne, médiane, étendue

 *exo 3 ; exo 4 ; exo 5 ; exo 6 ; 19 p 165 ; 53 p 172 ; exo 7*