# Probabilités et statistiques-séance 3

## Licence 1-semestre 1

Allegret Audrey

Maître de Conférences - Université de Toulon, LEAD

15 septembre 2021

# Plan

## Interprétation des hauteurs

La grande différence entre les diagrammes en bâtons et les histogrammes est que dans les premiers ni est représenté en hauteur tandis que, dans les seconds, il est représenté en surface.

## Interprétation des hauteurs

La grande différence entre les diagrammes en bâtons et les histogrammes est que dans les premiers ni est représenté en hauteur tandis que, dans les seconds, il est représenté en surface. Quelle est alors la signification de la hauteur dans un histogramme ? On a :

$$n_i = a_i * h_i$$

$\Rightarrow \mathrm{h}_i = \dfrac{n_i}{a_i}$ Le rapport $d_i = \dfrac{n_i}{a_i}$ est la densité de classe $c_i$.

# Histogrammes

## Interprétation des hauteurs

La grande différence entre les diagrammes en bâtons et les histogrammes est que dans les premiers ni est représenté en hauteur tandis que, dans les seconds, il est représenté en surface. Quelle est alors la signification de la hauteur dans un histogramme ? On a :

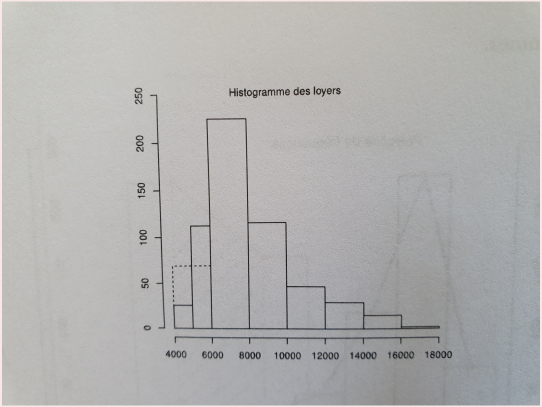
$$n_i = a_i * h_i$$

$\Rightarrow \mathrm{h}_i = \dfrac{n_i}{a_i}$ Le rapport $d_i = \dfrac{n_i}{a_i}$ est la densité de classe $c_i$.
Donc lorsqu'un rectangle est plus haut qu'un autre, c'est que la densité de son intervalle est plus grande, autrement dit qu'il comporte plus de données à amplitude égale.

**Interprétation des hauteurs**

Les histogrammes appréciés en hauteur donnent un aperçu de la densité de répartition des données.

## Interprétation des hauteurs



# Histogrammes

## Interprétation des hauteurs

Le rectangle en pointillés représente la fusion des deux premières classes.

## Interprétation des hauteurs

Le rectangle en pointillés représente la fusion des deux premières classes.
Leur effectif cumulé est de $13 + 56 = 69$ et le rectangle a donc une hauteur de 69 pour une amplitude de 2000.

# Histogrammes

## Interprétation des hauteurs

Le rectangle en pointillés représente la fusion des deux premières classes.
Leur effectif cumulé est de $13 + 56 = 69$ et le rectangle a donc une hauteur de 69 pour une amplitude de 2000.
Précédemment on avait deux rectangles de hauteurs respectives $2*13 = 26$ et $2*56 = 112$.

## Polygônes de fréquence

On obtient le polygône de fréquence en joignant, par une ligne polygonale, les points situés au milieu des arêtes supérieures des rectangles.

## Polygônes de fréquence

On obtient le polygône de fréquence en joignant, par une ligne polygonale, les points situés au milieu des arêtes supérieures des rectangles.
Ces graphiques permettent de visualiser les densités au moyen d'une ligne continue plutôt que par des paliers.

## Polygônes de fréquence

On obtient le polygône de fréquence en joignant, par une ligne polygonale, les points situés au milieu des arêtes supérieures des rectangles.
Ces graphiques permettent de visualiser les densités au moyen d'une ligne continue plutôt que par des paliers.
L'effet obtenu est de lisser les créneaux des histogrammes.

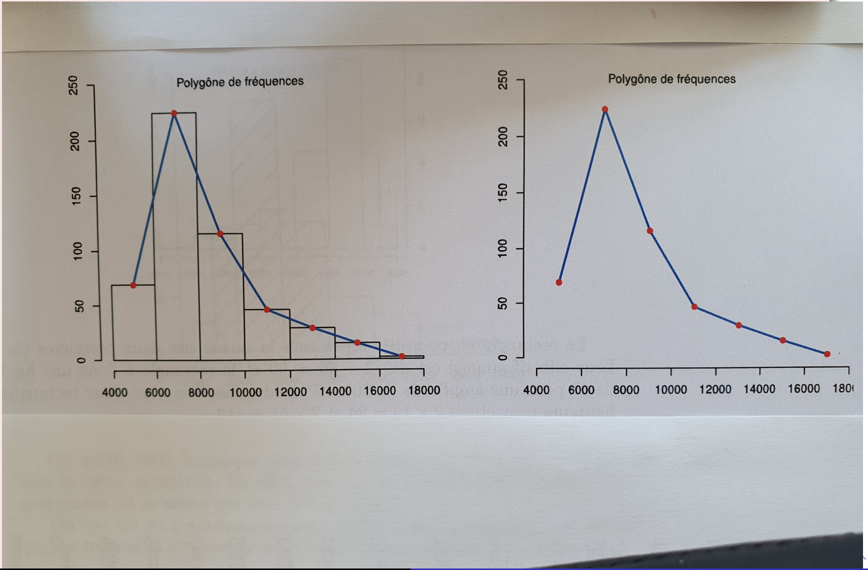
## Polygônes de fréquence

On obtient le polygône de fréquence en joignant, par une ligne polygonale, les points situés au milieu des arêtes supérieures des rectangles.
Ces graphiques permettent de visualiser les densités au moyen d'une ligne continue plutôt que par des paliers.
L'effet obtenu est de lisser les créneaux des histogrammes.
On utilise cette technique pour des histogrammes dont les rectangles ont tous la même amplitude.

## Polygônes de fréquence

On obtient le polygône de fréquence en joignant, par une ligne polygonale, les points situés au milieu des arêtes supérieures des rectangles.
Ces graphiques permettent de visualiser les densités au moyen d'une ligne continue plutôt que par des paliers.
L'effet obtenu est de lisser les créneaux des histogrammes.
On utilise cette technique pour des histogrammes dont les rectangles ont tous la même amplitude.
En effet, dans ce cas, la surface située sous la courbe polygonale est la même que celle des rectangles.

## Polygônes de fréquence



## Diagrammes de dispersion

Les diagrammes de dispersion servent à représenter les corrélations qui peuvent exister entre des observations portant sur deux variables différentes.

## Diagrammes de dispersion

Les diagrammes de dispersion servent à représenter les corrélations qui peuvent exister entre des observations portant sur deux variables différentes.
Si x et y sont les deux variables observées, pour chaque observation $O_i$, on place le point de coordonnées $(x_i; y_i)$.

## Diagrammes de dispersion

Les diagrammes de dispersion servent à représenter les corrélations qui peuvent exister entre des observations portant sur deux variables différentes.
Si x et y sont les deux variables observées, pour chaque observation $O_i$, on place le point de coordonnées $(x_i; y_i)$.
L'ensemble des points obtenus s'appelle un nuage.

## Diagrammes de dispersion

Les diagrammes de dispersion servent à représenter les corrélations qui peuvent exister entre des observations portant sur deux variables différentes.
Si x et y sont les deux variables observées, pour chaque observation $O_i$, on place le point de coordonnées $(x_i; y_i)$.
L'ensemble des points obtenus s'appelle un nuage.
On voit que les points obtenus sont uniformément répartis dans le carré.

## Diagrammes de dispersion

Les diagrammes de dispersion servent à représenter les corrélations qui peuvent exister entre des observations portant sur deux variables différentes.
Si x et y sont les deux variables observées, pour chaque observation $O_i$, on place le point de coordonnées $(x_i; y_i)$.
L'ensemble des points obtenus s'appelle un nuage.
On voit que les points obtenus sont uniformément répartis dans le carré.
Cet exemple est typique de **l'absence de corrélation** entre les variables x et y.

## Diagrammes de dispersion

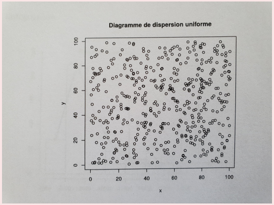
L'ensemble des points obtenus s'appelle un nuage.

## Diagrammes de dispersion

L'ensemble des points obtenus s'appelle un nuage.
Le graphique qui suit est un diagramme de dispersion correspondant à deux variables x et y **distribuées uniformément**.

## Diagrammes de dispersion

L'ensemble des points obtenus s'appelle un nuage.
Le graphique qui suit est un diagramme de dispersion correspondant à deux variables x et y **distribuées uniformément**.
Cela signifie simplement que les valeurs observées pour chacune des deux variables sont équiréparties sur l'intervalle où elles sont définies.



## Diagrammes de dispersion

Les deux graphiques qui suivent sont des diagrammes de dispersion correspondant à deux variables x et y **distribuées**.

## Diagrammes de dispersion

Les deux graphiques qui suivent sont des diagrammes de dispersion correspondant à deux variables x et y **distribuées**.
Cette notion sera vue avec précision par la suite, mais ici cela signifie simplement que les observations sont massées autour d'une valeur centrale et qu'elles se raréfient quand on s'en éloigne.

## Diagrammes de dispersion

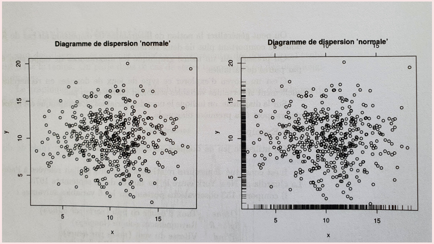
Les deux graphiques qui suivent sont des diagrammes de dispersion correspondant à deux variables x et y **distribuées**.
Cette notion sera vue avec précision par la suite, mais ici cela signifie simplement que les observations sont massées autour d'une valeur centrale et qu'elles se raréfient quand on s'en éloigne.
Le deuxième graphique ajoute justement des petits traits (le long des axes) qui matérialisent la répartition des x et des y.

## Diagrammes de dispersion

On voit que les points obtenus sont uniformément répartis autour d'un point central.

## Diagrammes de dispersion

On voit que les points obtenus sont uniformément répartis autour d'un point central.
Cet exemple est typique de l'absence de corrélation entre les variables x et y.

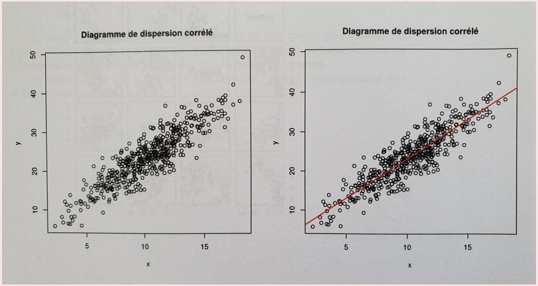


## Diagrammes de dispersion

L'exemple qui suit est un diagramme qui suggère une **corrélation positive** entre les x et les y.

## Diagrammes de dispersion

L'exemple qui suit est un diagramme qui suggère une **corrélation positive** entre les x et les y. Le nuage de points semble orienté dans une direction particulière.



## Graphique de séries temporelles

Ce sont des diagrammes qui sont utilisés pour représenter des données qui évoluent dans le temps.

## Graphique de séries temporelles

Ce sont des diagrammes qui sont utilisés pour représenter des données qui évoluent dans le temps.
On parle dans ce cas de **séries temporelles**.
3-16.jpg

# Exercice

## Application sous excel

A l'aide des données figurant dans votre fichier excel :

1. Faire un diagramme circulaire représentant les proportions des diverses catégories.
2. Faire un diagramme en bâton représentant les répartitions.
3. Faire un diagramme en bâton représentant les répartitions en distinguant les hommes et les femmes.

### Introduction

Les indicateurs de position fournissent des renseignements sur des variables aussi bien qualitatives que quantitatives.

## Introduction

Les indicateurs de position fournissent des renseignements sur des variables aussi bien qualitatives que quantitatives.
Les caractéristiques de position sont de deux types :

- Celles relatives aux effectifs : mode et classe modale.

### Introduction

Les indicateurs de position fournissent des renseignements sur des variables aussi bien qualitatives que quantitatives.
Les caractéristiques de position sont de deux types :

- Celles relatives aux effectifs : mode et classe modale.
- D'autres sont relatives au rang occupé par les observations les unes par rapport aux autres plutôt qu'à leur valeur.

## Introduction

Les indicateurs de position fournissent des renseignements sur des variables aussi bien qualitatives que quantitatives.
Les caractéristiques de position sont de deux types :

- Celles relatives aux effectifs : mode et classe modale.
- D'autres sont relatives au rang occupé par les observations les unes par rapport aux autres plutôt qu'à leur valeur.

**Remarque** : dans le cas d'une variable qualitative, il faut qu'elle soit ordinale, c'est-à-dire qu'on puisse ordonner les valeurs du caractère, si on veut pouvoir parler de rang.

### Mode

Le mode concerne les variables qualitatives ou quantitatives discrètes. Dans ce cas, on dresse la table des effectifs qui dénombre les observations correspondant à chaque modalité.

## Mode

Le mode concerne les variables qualitatives ou quantitatives discrètes. Dans ce cas, on dresse la table des effectifs qui dénombre les observations correspondant à chaque modalité.
Par définition, **le mode est la valeur (ou la modalité) de la variable qui a l'effectif le plus élevé**.

### Mode

Le mode concerne les variables qualitatives ou quantitatives discrètes. Dans ce cas, on dresse la table des effectifs qui dénombre les observations correspondant à chaque modalité.
Par définition, **le mode est la valeur (ou la modalité) de la variable qui a l'effectif le plus élevé**.
Sur un diagramme en bâtons, c'est la modalité qui correspond au bâton le plus haut.

## Mode : exemple I

Une enquête de satisfaction a attribué une note entre 1 et 10 pour évaluer la qualité d'un service. Les résultats sont les suivants :

## Mode : exemple I

Une enquête de satisfaction a attribué une note entre 1 et 10 pour évaluer la qualité d'un service. Les résultats sont les suivants :

| Notes | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Effectifs | 5 | 3 | 4 | 7 | 8 | 10 | 15 | 11 | 6 | 6 |

Il s'agit d'une variable quantitative discrète.

# Caractéristiques de position

## Mode : exemple I

Une enquête de satisfaction a attribué une note entre 1 et 10 pour évaluer la qualité d'un service. Les résultats sont les suivants :

| Notes | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Effectifs | 5 | 3 | 4 | 7 | 8 | 10 | 15 | 11 | 6 | 6 |

Il s'agit d'une variable quantitative discrète.

Le mode est 7.

## Mode : exemple I

Une enquête de satisfaction a attribué une note entre 1 et 10 pour évaluer la qualité d'un service. Les résultats sont les suivants :

| Notes | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Effectifs | 5 | 3 | 4 | 7 | 8 | 10 | 15 | 11 | 6 | 6 |

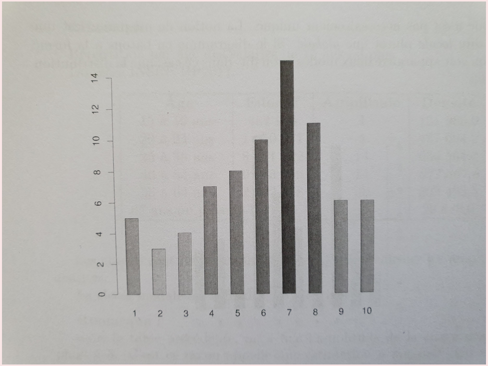
Il s'agit d'une variable quantitative discrète.

Le mode est 7.

Le diagramme en bâtons ci-dessous fait clairement apparaître le mode.

# Caractéristiques de position

## Calcul graphique du mode



## Mode : exemple 2

Prenons l'exemple de type de ménage selon la structure familiale.

| Type de ménage | en milliers |
|---|---|
| Homme seul | 4032.2 |
| Femme seule | 5529.5 |
| Couple sans enfant | 7250.4 |
| Couple avec enfants | 7435.6 |
| Famille monoparentale | 2345.2 |

## Mode : exemple 2

Prenons l'exemple de type de ménage selon la structure familiale.

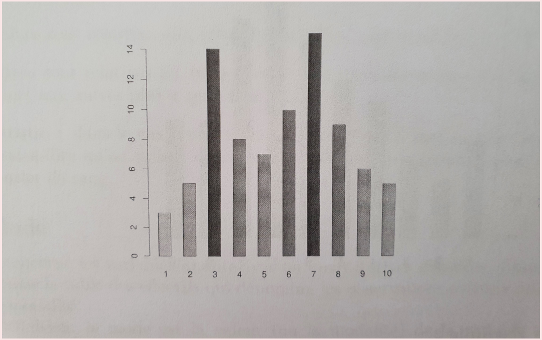
| Type de ménage | en milliers |
|---|---|
| Homme seul | 4032.2 |
| Femme seule | 5529.5 |
| Couple sans enfant | 7250.4 |
| Couple avec enfants | 7435.6 |
| Famille monoparentale | 2345.2 |

Ici les données sont en effectifs. Il s'agit d'une variable qualitative.

Le **mode est "couple avec enfant(s)"**.

## Mode

Le mode n'est pas nécessairement unique. Si le diagramme en bâtons a la forme suivante, on voit apparaître deux modes et on dit, dans ce cas, que la distribution est bimodale.



## Classe modale

La classe modale concerne les variables quantitatives continues.

# Caractéristiques de position

## Classe modale

La classe modale concerne les variables quantitatives continues. Dans ce cas, on regroupe les données en classes et on dresse la table des effectifs qui dénombre les observations entrant dans chacune des classes.

## Classe modale

La classe modale concerne les variables quantitatives continues. Dans ce cas, on regroupe les données en classes et on dresse la table des effectifs qui dénombre les observations entrant dans chacune des classes.
On distingue alors deux cas :

- si les classes sont d'amplitude égale, la classe modale est celle qui a le plus grand effectif.

## Classe modale

La classe modale concerne les variables quantitatives continues. Dans ce cas, on regroupe les données en classes et on dresse la table des effectifs qui dénombre les observations entrant dans chacune des classes.
On distingue alors deux cas :

- si les classes sont d'amplitude égale, la classe modale est celle qui a le plus grand effectif.
- si les classes ne sont pas d'amplitude égale, on ramène tout à une amplitude commune en divisant les fréquences par la longueur des intervalles. On calcule donc les densités de chaque classe :

## Classe modale

$$d_i = \frac{n_i}{a_i}$$

Par définition, la classe modale est celle de plus forte densité.

## Classe modale

$$d_i = \frac{n_i}{a_i}$$

Par définition, la classe modale est celle de plus forte densité. Graphiquement, on la repère sur un histogramme comme étant celle dont le rectangle est le plus haut (voir ci-dessous).

## Classe modale : exemple I

Une entreprise s'intéresse à la distance parcourue par les employés entre le domicile et le lieu de travail. On a noté les effectifs suivants :

## Classe modale : exemple I

Une entreprise s'intéresse à la distance parcourue par les employés entre le domicile et le lieu de travail. On a noté les effectifs suivants :

| km | [0 ;10[ | [10 ;20[ | [20 ;30[ | [30 ;40[ | [40 ;50[ | [50 ;60[ | [60 ;70[ |
|---|---|---|---|---|---|---|---|
| Effectifs | 14 | 23 | 32 | 21 | 12 | 8 | 4 |

Ici toutes les classes sont de même amplitude. La classe modale est [20 ;30[.

## Classe modale : exemple II

Soit la répartition par âge des salariés de 15 ans ou plus :

## Classe modale : exemple II

Soit la répartition par âge des salariés de 15 ans ou plus :

| âge | effectif | amplitude | densité |
|---|---|---|---|
| [15 ;19[ | 487944 | 4 | 121986 |
| [20 ;24[ | 1950777 | 4 | 487694.2 |
| [25 ;39[ | 8911762 | 14 | 636554.4 |
| [40 ;54[ | 9483149 | 14 | 677367.8 |
| [55 ;64[ | 2722458 | 9 | 302495.3 |
| [65 ;70[ | 149400 | 5 | 29880 |

Les **classes sont de longueur inégale**.

## Classe modale : exemple II

Soit la répartition par âge des salariés de 15 ans ou plus :

| âge | effectif | amplitude | densité |
|---|---|---|---|
| [15 ;19[ | 487944 | 4 | 121986 |
| [20 ;24[ | 1950777 | 4 | 487694.2 |
| [25 ;39[ | 8911762 | 14 | 636554.4 |
| [40 ;54[ | 9483149 | 14 | 677367.8 |
| [55 ;64[ | 2722458 | 9 | 302495.3 |
| [65 ;70[ | 149400 | 5 | 29880 |

Les **classes sont de longueur inégale**.

On doit calculer les amplitudes et les densités. La classe modale est celle des “40 à 54 ans”.

### Médiane

La notion de médiane concerne les variables quantitatives.

## Médiane

La notion de médiane concerne les variables quantitatives.
La médiane est une quantité qui partage les observations en deux groupes de même taille.

**Médiane**

La notion de médiane concerne les variables quantitatives.
La médiane est une quantité qui partage les observations en deux groupes de même taille.
C'est donc une valeur M telle qu'il y ait 50% des observations pour lesquelles le caractère observé X est inférieur à M et 50% des observations pour lesquelles le caractère observé X est supérieur à M.

## Médiane : exemple I

On a relevé les notes de 9 étudiants à un examen :

## Médiane : exemple I

On a relevé les notes de 9 étudiants à un examen :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 11.5 | 7 | 16 | 14 | 8.5 | 10.5 | 13 | 11 | 5 |

## Médiane : exemple I

On a relevé les notes de 9 étudiants à un examen :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 11.5 | 7 | 16 | 14 | 8.5 | 10.5 | 13 | 11 | 5 |

Pour trouver la médiane, il faut commencer par ordonner les notes :

### Médiane : exemple I

On a relevé les notes de 9 étudiants à un examen :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 11.5 | 7 | 16 | 14 | 8.5 | 10.5 | 13 | 11 | 5 |

Pour trouver la médiane, il faut commencer par ordonner les notes :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 5 | 7 | 8.5 | 10.5 | 11 | 11.5 | 13 | 14 | 16 |

La valeur $M = 11$ est la médiane car elle sépare les données en deux groupes de même taille.

## Médiane : exemple II

Prenons le cas d'une variable continue.

## Médiane : exemple II

Prenons le cas d'une variable continue.
Considérons les Exploitations agricoles selon la superficie agricole utilisée (SAU) en 2010.

## Médiane : exemple II

Prenons le cas d'une variable continue.
Considérons les Exploitations agricoles selon la superficie agricole utilisée (SAU) en 2010.

| SAU | Effectifs | Proportions | Prop. cumulées |
|---|---|---|---|
| Moins de 20 ha | 235,4 | 45.74 | *45.74* |
| De 20 à moins de 50 ha | 88,4 | 17.18 | *62.92* |
| De 50 à moins de 100 ha | 97,6 | 18.96 | 81.88 |
| De 100 à moins de 200 ha | 72,7 | 14.12 | 96.00 |
| 200 ha ou plus | 20,6 | 4.00 | 100 |
| **Total** | 514,7 | | |

## Médiane : exemple II

On cherche où se situe la proportion cumulée de 50%.

## Médiane : exemple II

On cherche où se situe la proportion cumulée de 50%.
D'après la tableau, c'est entre 45.74% et 62.92%.

## Médiane : exemple II

On cherche où se situe la proportion cumulée de 50%.
D'après la tableau, c'est entre 45.74% et 62.92%.
Il faut faire une interpolation linéaire pour déterminer la médiane.

## Médiane : exemple II

On cherche où se situe la proportion cumulée de 50%.
D'après la tableau, c'est entre 45.74% et 62.92%.
Il faut faire une interpolation linéaire pour déterminer la médiane.
L'interpolation linéaire consiste à chercher la valeur M qui soit par rapport à 20 et 50 comme la valeur 50% par rapport à 45.74% et 62.92% :

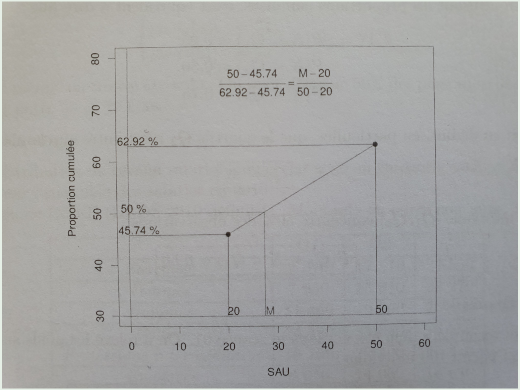
## Médiane : exemple II

On cherche où se situe la proportion cumulée de 50%.
D'après la tableau, c'est entre 45.74% et 62.92%.
Il faut faire une interpolation linéaire pour déterminer la médiane.
L'interpolation linéaire consiste à chercher la valeur M qui soit par rapport à 20 et 50 comme la valeur 50% par rapport à 45.74% et 62.92% :

| SAU | Fréquences cumulées |
|---|---|
| 20 ha | 45.74% |
| M ? | M ? |
| 50 ha | 62.92% |

Graphiquement, on représente le problème de la manière suivante :

## Médiane : exemple II



# Caractéristiques de position

## Médiane : exemple III

Numériquement :

$$\frac{50-45.74}{62.92-45.74} = \frac{M-20}{50-20}$$
$$\frac{4.26}{17.18} = \frac{M-20}{30}$$
$$M = 20 + \frac{4.26*30}{17.18} = 27.44$$

## Définition des quantiles

Les quantiles sont des valeurs permettant de partager les observations ordonnées d'une série en sous-groupes contenant le même nombre de données (aux erreurs d'arrondis près).

## Définition des quantiles

Les quantiles sont des valeurs permettant de partager les observations ordonnées d'une série en sous-groupes contenant le même nombre de données (aux erreurs d'arrondis près).
On utilise le plus souvent, les trois quantiles suivants : les quartiles, les déciles et les centiles.

### quartiles

Les **quartiles** sont trois quantités qui partagent les observations en quatre groupes de même taille.

## quartiles

Les **quartiles** sont trois quantités qui partagent les observations en quatre groupes de même taille.
Ce sont des valeurs habituellement notées $Q_1$, $Q_2$ et $Q_3$, telles qu'il y ait 25% des observations pour lesquelles le caractère observé X soit compris dans les intervalles qu'elles délimitent. On a :
$P(X < Q_1) = 0.25$

**quartiles**

Les **quartiles** sont trois quantités qui partagent les observations en quatre groupes de même taille.
Ce sont des valeurs habituellement notées $Q_1$, $Q_2$ et $Q_3$, telles qu'il y ait 25% des observations pour lesquelles le caractère observé X soit compris dans les intervalles qu'elles délimitent. On a :
$P(X < Q_1) = 0.25$
$P(Q_1 < X < Q_2) = 0.25$

## quartiles

Les **quartiles** sont trois quantités qui partagent les observations en quatre groupes de même taille.
Ce sont des valeurs habituellement notées $Q_1$, $Q_2$ et $Q_3$, telles qu'il y ait 25% des observations pour lesquelles le caractère observé X soit compris dans les intervalles qu'elles délimitent. On a :

$P(X < Q_1) = 0.25$
$P(Q_1 < X < Q_2) = 0.25$
$P(Q_2 < X < Q_3) = 0.25$

# Caractéristiques de position

## quartiles

Les **quartiles** sont trois quantités qui partagent les observations en quatre groupes de même taille.
Ce sont des valeurs habituellement notées $Q_1$, $Q_2$ et $Q_3$, telles qu'il y ait 25% des observations pour lesquelles le caractère observé X soit compris dans les intervalles qu'elles délimitent. On a :

$P(X < Q_1) = 0.25$
$P(Q_1 < X < Q_2) = 0.25$
$P(Q_2 < X < Q_3) = 0.25$
$P(X > Q_3) = 0.25$

## quartiles

En utilisant les fréquences cumulées, c'est équivalent à :
$P(X < Q_1) = 0.25$

# Caractéristiques de position

## quartiles

En utilisant les fréquences cumulées, c'est équivalent à :

$P(X < Q_1) = 0.25$

$P(X < Q_2) = 0.5$

# Caractéristiques de position

## quartiles

En utilisant les fréquences cumulées, c'est équivalent à :
$P(X < Q_1) = 0.25$
$P(X < Q_2) = 0.5$
$P(X < Q_3) = 0.75$

### quartiles

En utilisant les fréquences cumulées, c'est équivalent à :
$P(X < Q_1) = 0.25$
$P(X < Q_2) = 0.5$
$P(X < Q_3) = 0.75$
On en déduit un cas particulier, que le quartile $Q_2$ n'est autre que la médiane.

## quartiles

En utilisant les fréquences cumulées, c'est équivalent à :
$P(X < Q_1) = 0.25$
$P(X < Q_2) = 0.5$
$P(X < Q_3) = 0.75$
On en déduit un cas particulier, que le quartile $Q_2$ n'est autre que la médiane.
A noter également que l'intervalle $[Q_1, Q_3]$ concentre 50% des observations.

# Caractéristiques de position

## Exemple

Soit les données suivantes :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 64 | 85 | 79 | 84 | 68 | 74 | 94 | 75 | 64 | 65 |
| 72 | 74 | 78 | 69 | 67 | 64 | 70 | 63 | 69 | 82 |
| 62 | 64 | 71 | 74 | 77 | 73 | 77 | 76 | 82 | 82 |
| 86 | 48 | 50 | 69 | 76 | 59 | 70 | 61 | 55 | 77 |
| 73 | 81 | 76 | 56 | 63 | 84 | 63 | 57 | 76 | 86 |
| 62 | 70 | 69 | 66 | 63 | 90 | 72 | 73 | 73 | 76 |
| 75 | 70 | 68 | 66 | 74 | 66 | 52 | 66 | 81 | 57 |
| 77 | 79 | 55 | 69 | 78 | 60 | 85 | 70 | 67 | 64 |
| 76 | 78 | 65 | 81 | 69 | 76 | 72 | 71 | 74 | 58 |
| 67 | 76 | 74 | 78 | 79 | 69 | 92 | 64 | 73 | 65 |

### Exemple

On doit d'abord ordonner les poids :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 48 | 50 | 52 | 55 | 55 | 56 | 57 | 57 | 58 | 59 |
| 60 | 61 | 62 | 62 | 63 | 63 | 63 | 63 | 64 | 64 |
| 64 | 64 | 64 | 64 | *65* | *65* | 65 | 66 | 66 | 66 |
| 66 | 67 | 67 | 67 | 68 | 68 | 69 | 69 | 69 | 69 |
| 69 | 69 | 69 | 70 | 70 | 70 | 70 | 70 | 71 | *71* |
| *72* | 72 | 72 | 73 | 73 | 73 | 73 | 73 | 74 | 74 |
| 74 | 74 | 74 | 74 | 75 | 75 | 76 | 76 | 76 | 76 |
| 76 | 76 | 76 | 76 | *77* | *77* | 77 | 77 | 78 | 78 |
| 78 | 78 | 79 | 79 | 79 | 81 | 81 | 81 | 82 | 82 |
| 82 | 84 | 84 | 85 | 85 | 86 | 86 | 90 | 92 | 94 |

On a fait apparaître les valeurs situées en positions 25-26, 50-51 et 75-76.

# Caractéristiques de position

## Exemple

On doit d'abord ordonner les poids :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 48 | 50 | 52 | 55 | 55 | 56 | 57 | 57 | 58 | 59 |
| 60 | 61 | 62 | 62 | 63 | 63 | 63 | 63 | 64 | 64 |
| 64 | 64 | 64 | 64 | *65* | *65* | 65 | 66 | 66 | 66 |
| 66 | 67 | 67 | 67 | 68 | 68 | 69 | 69 | 69 | 69 |
| 69 | 69 | 69 | 70 | 70 | 70 | 70 | 70 | 71 | *71* |
| *72* | 72 | 72 | 73 | 73 | 73 | 73 | 73 | 74 | 74 |
| 74 | 74 | 74 | 74 | 75 | 75 | 76 | 76 | 76 | 76 |
| 76 | 76 | 76 | 76 | *77* | *77* | 77 | 77 | 78 | 78 |
| 78 | 78 | 79 | 79 | 79 | 81 | 81 | 81 | 82 | 82 |
| 82 | 84 | 84 | 85 | 85 | 86 | 86 | 90 | 92 | 94 |

On a fait apparaître les valeurs situées en positions 25-26, 50-51 et 75-76.
Ce sont des intervalles quartiles. Il faut en prendre le milieu.

**Exemple**

Soit :

$$\begin{cases} poids[25] = 65 \\ poids[26] = 65 \end{cases}$$

**Exemple**

$\Rightarrow Q_1 = 65.$

## Exemple

$\Rightarrow Q_1 = 65.$

Pour le deuxième quartile nous procédons de la même manière :

$$\begin{cases} poids[50] = 71 \\ poids[51] = 72 \end{cases}$$

$\Rightarrow Q_2 = 71.5.$

**Exemple**

$\Rightarrow Q_1 = 65.$

Pour le deuxième quartile nous procédons de la même manière :

$$\begin{cases} poids[50] = 71 \\ poids[51] = 72 \end{cases}$$

$\Rightarrow Q_2 = 71.5.$

De même, on trouve $Q_3 = 77$.

## Exemple

$\Rightarrow Q_1 = 65$.

Pour le deuxième quartile nous procédons de la même manière :

$$\begin{cases} poids[50] = 71 \\ poids[51] = 72 \end{cases}$$

$\Rightarrow Q_2 = 71.5$.

De même, on trouve $Q_3 = 77$.

On peut dire que 50% des personnes observées pèsent entre 65 et 77 kilos.