

# Description des données

Thibaut FABACHER

GMRC

# Plan

1. Correction du devoir

# Plan

1. Correction du devoir
2. Rappels

# Plan

1. Correction du devoir
2. Rappels
3. Définitions

# Plan

1. Correction du devoir
2. Rappels
3. Définitions
4. Statistiques descriptives

# Correction du devoir

Éléments importants

Encodages de caractères



A	Ω	語	𐄌
00000041	000003A9	00008A9E	00010384

UTF-32

A	Ω	語	𐄌
0041	03A9	8A9E	D800   DF84

UTF-16

A	Ω	語	𐄌
41	CE   A9	E8   AA   9E	F0   90   8E   84

UTF-8

## Application de fonction

```
# Executer une fonction
```

```
fonction(arg1 = ... , arg2 = ... , arg3 = ....)
```

```
# L'attribuer à un objet la sortie de la fonction
```

# Rappels sur les statistiques



## Les statistiques au lycée

- Moyennes, probabilités

Exemple : Notes d'un élève

Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

## Les statistiques au lycée

- ▶ Moyennes, probabilités
- ▶ Espérance et variance ?

### Exemple : Notes d'un élève

#### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

## Les statistiques au lycée

- ▶ Moyennes, probabilités
- ▶ Espérance et variance ?
- ▶ Loi Bernoulli / binomiale ?

### Exemple : Notes d'un élève

#### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

## Les statistiques au lycée

- ▶ Moyennes, probabilités
- ▶ Espérance et variance ?
- ▶ Loi Bernoulli / binomiale ?

### Exemple : Notes d'un élève

- ▶ Notes d'un élève, relevées sur une année

### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

## Les statistiques au lycée

- Moyennes, probabilités
- Espérance et variance ?
- Loi Bernoulli / binomiale ?

### Exemple : Notes d'un élève

- Notes d'un élève, relevées sur une année
- Contrôle de biologie

### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

## Les statistiques au lycée

- ▶ Moyennes, probabilités
- ▶ Espérance et variance ?
- ▶ Loi Bernoulli / binomiale ?

### Exemple : Notes d'un élève

- ▶ Notes d'un élève, relevées sur une année
- ▶ Contrôle de biologie
- ▶ 5 notes sur le trimestre

### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

## Les statistiques au lycée

- Moyennes, probabilités
- Espérance et variance ?
- Loi Bernoulli / binomiale ?

### Exemple : Notes d'un élève

- Notes d'un élève, relevées sur une année
- Contrôle de biologie
- 5 notes sur le trimestre

### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

- Moy :  $\frac{11+8+9+10+11}{5} = 9.8$

## Les statistiques au lycée

- ▶ Moyennes, probabilités
- ▶ Espérance et variance ?
- ▶ Loi Bernoulli / binomiale ?

### Exemple : Notes d'un élève

- ▶ Notes d'un élève, relevées sur une année
- ▶ Contrôle de biologie
- ▶ 5 notes sur le trimestre

### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

- ▶ Moy :  $\frac{11+8+9+10+11}{5} = 9.8$
- ▶ Bon indicateur d'un profil moyen



## Les statistiques au lycée

- ▶ Moyennes, probabilités
- ▶ Espérance et variance ?
- ▶ Loi Bernoulli / binomiale ?

### Exemple : Notes d'un élève

- ▶ Notes d'un élève, relevées sur une année
- ▶ Contrôle de biologie
- ▶ 5 notes sur le trimestre

### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

- ▶ Moy :  $\frac{11+8+9+10+11}{5} = 9.8$
- ▶ Bon indicateur d'un profil moyen
- ▶ Pas de notion de régularité

## Les statistiques au lycée

- ▶ Moyennes, probabilités
- ▶ Espérance et variance ?
- ▶ Loi Bernoulli / binomiale ?

### Exemple : Notes d'un élève

- ▶ Notes d'un élève, relevées sur une année
- ▶ Contrôle de biologie
- ▶ 5 notes sur le trimestre

### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

- ▶ Moy :  $\frac{11+8+9+10+11}{5} = 9.8$
- ▶ Bon indicateur d'un profil moyen
- ▶ Pas de notion de régularité

## Les statistiques au lycée

- ▶ Moyennes, probabilités
- ▶ Espérance et variance ?
- ▶ Loi Bernoulli / binomiale ?

### Exemple : Notes d'un élève

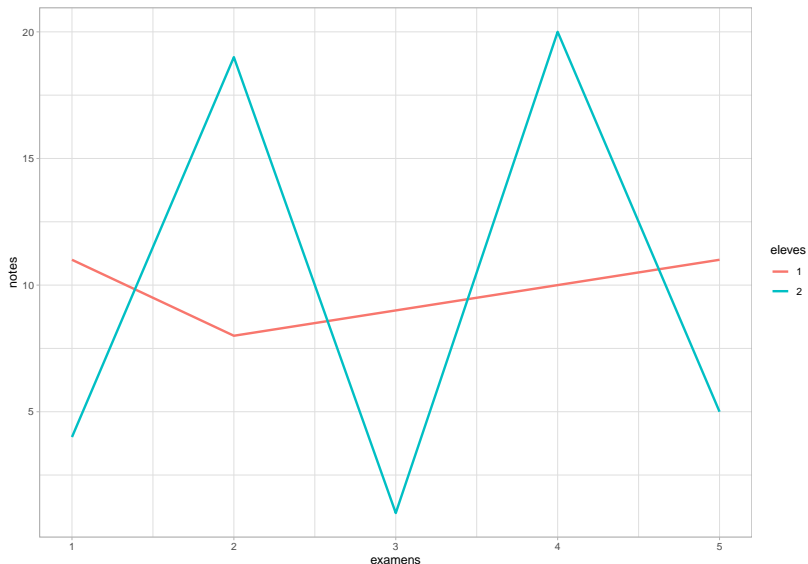
- ▶ Notes d'un élève, relevées sur une année
- ▶ Contrôle de biologie
- ▶ 5 notes sur le trimestre

### Résultats d'un élève

11	8	9	10	11
----	---	---	----	----

- ▶ Moy :  $\frac{11+8+9+10+11}{5} = 9.8$
- ▶ Bon indicateur d'un profil moyen
- ▶ Pas de notion de régularité

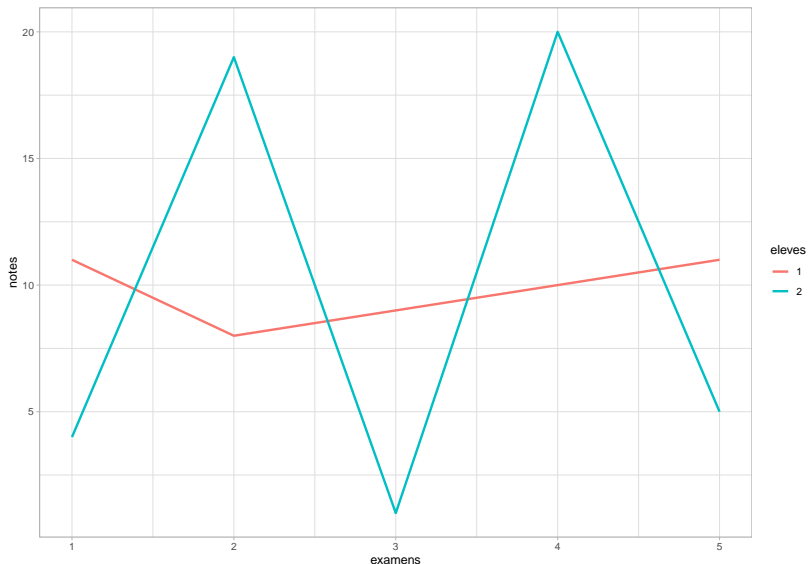
## Quelle différence ?



## La variabilité

► Phénomène de dispersion

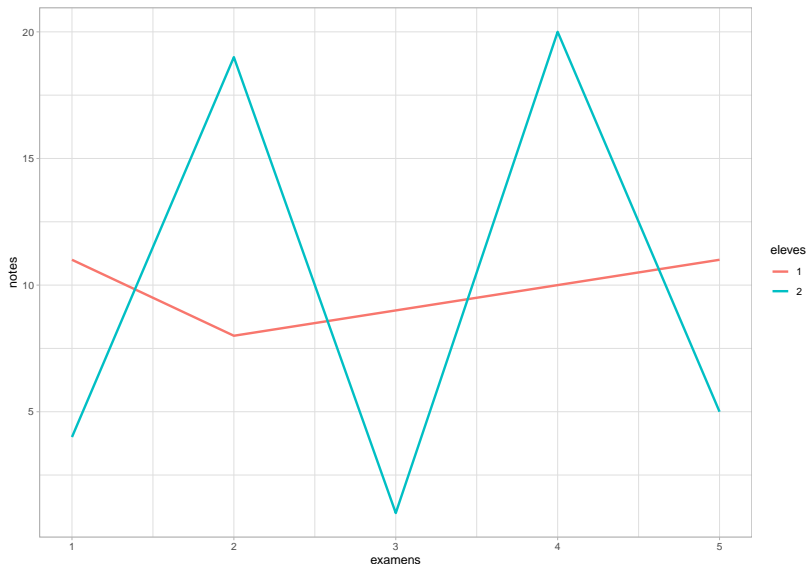
## Quelle différence ?



## La variabilité

► Phénomène de dispersion

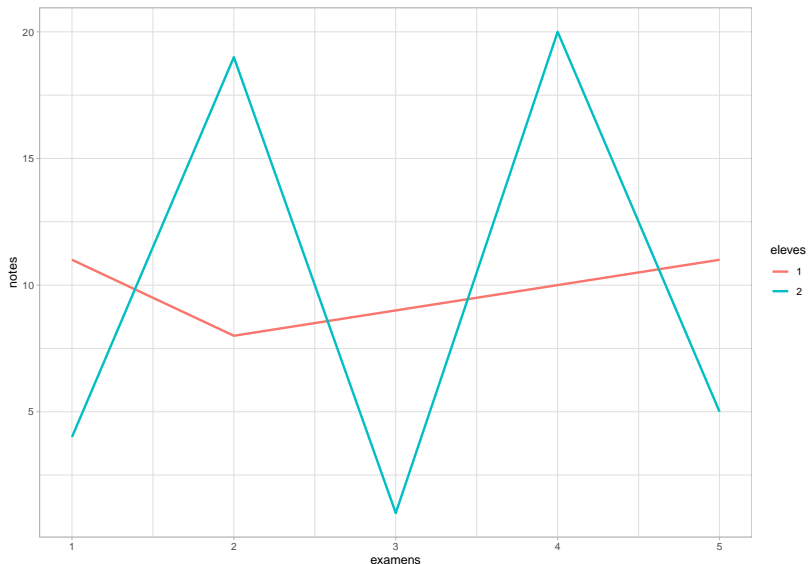
## Quelle différence ?



## La variabilité

► Phénomène de dispersion

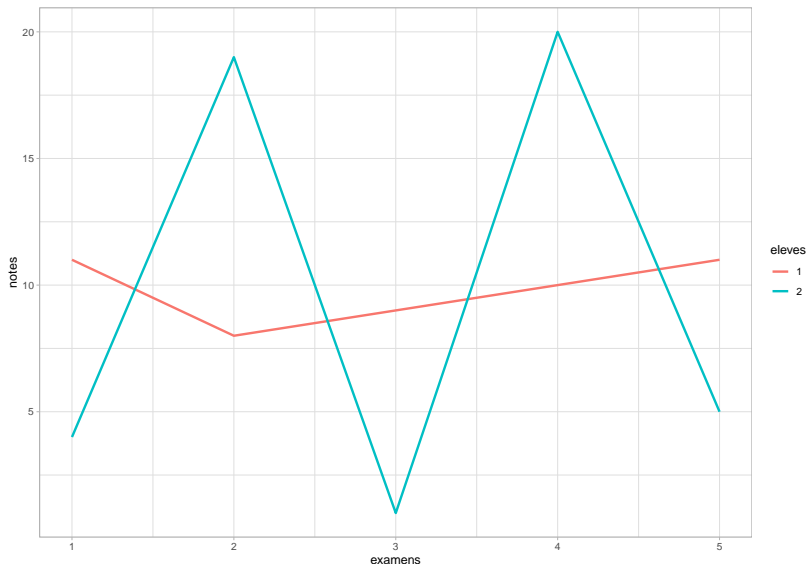
## Quelle différence ?



## La variabilité

► Phénomène de dispersion

## Quelle différence ?

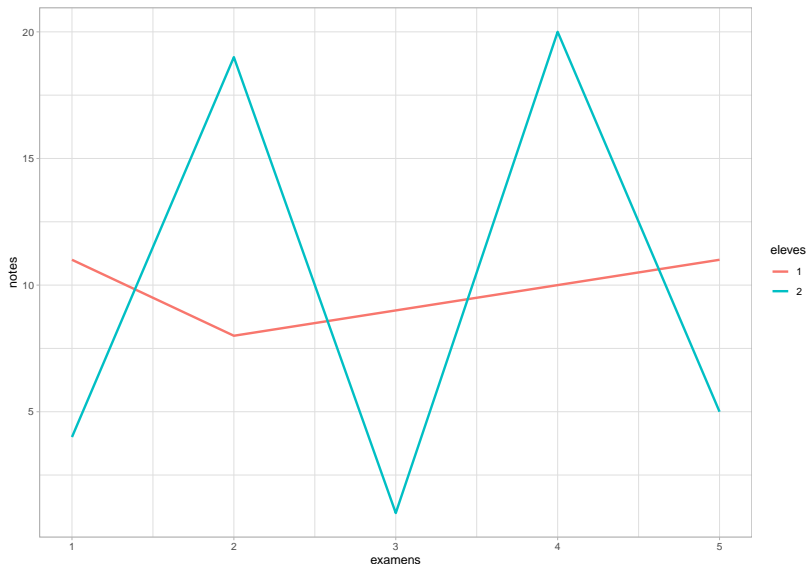


## La variabilité

► Phénomène de dispersion



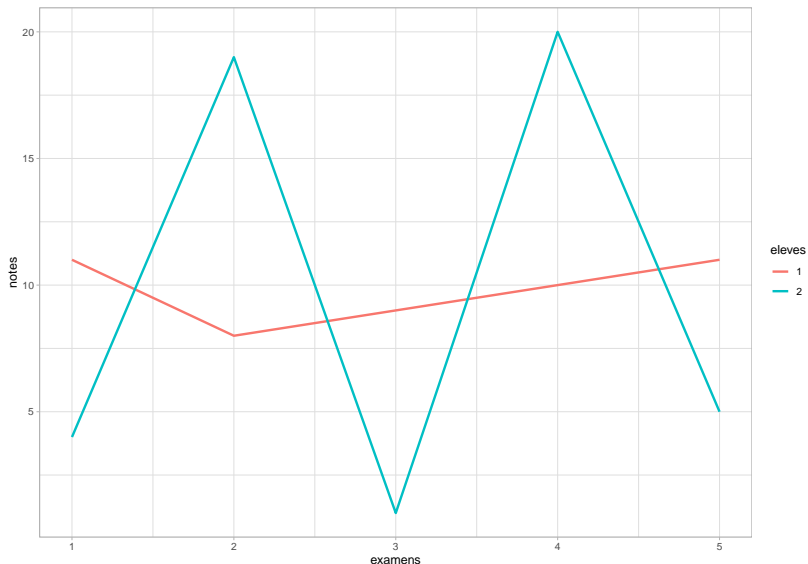
## Quelle différence ?



## La variabilité

► Phénomène de dispersion

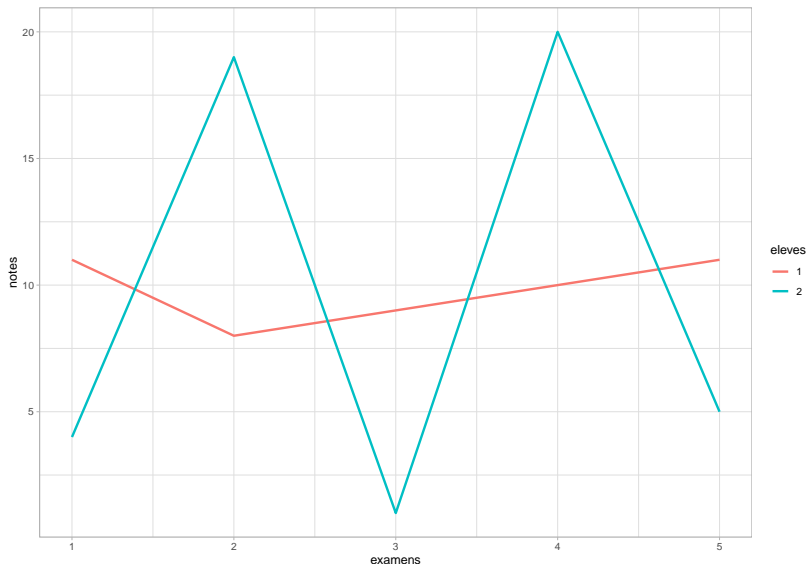
## Quelle différence ?



## La variabilité

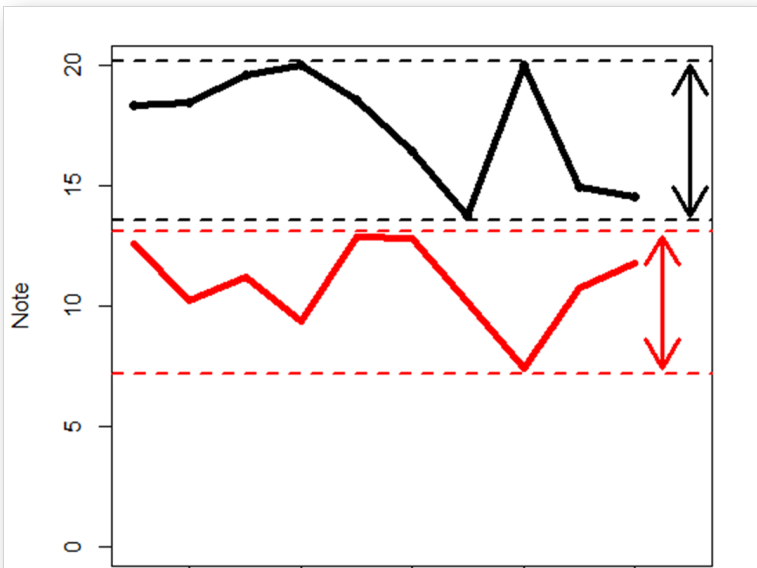
► Phénomène de dispersion

## Quelle différence ?



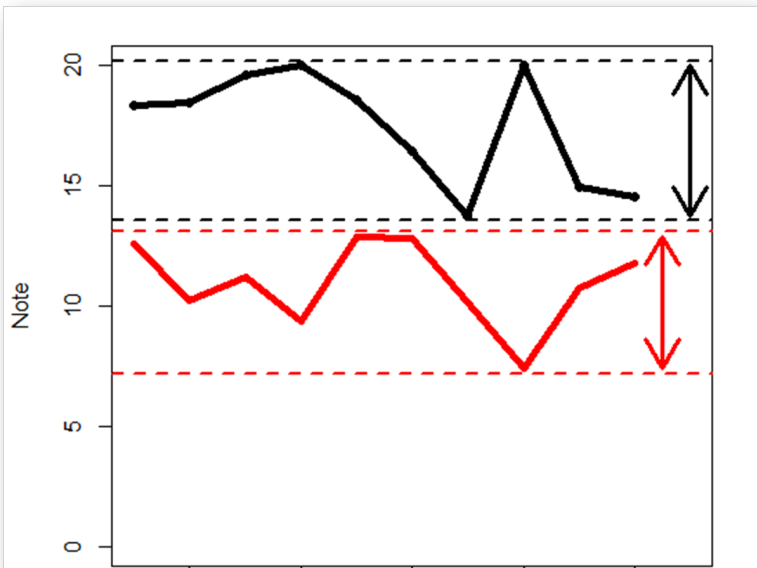
## La variabilité

► Phénomène de dispersion



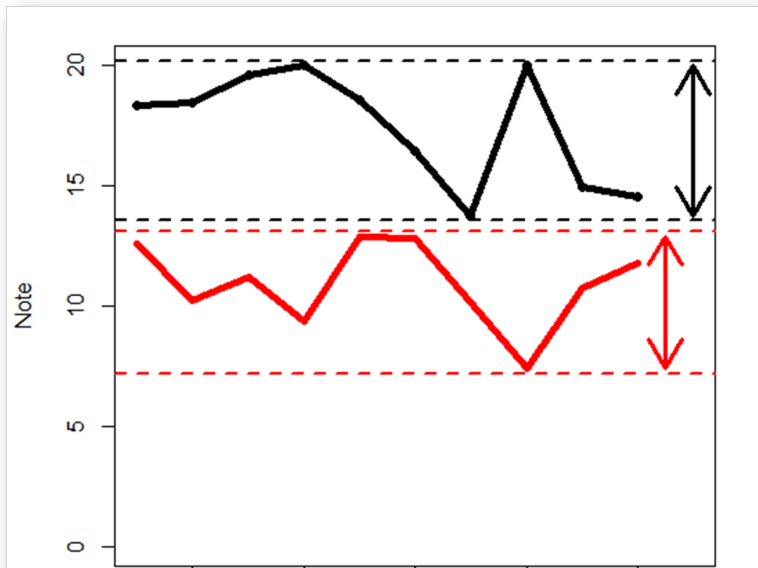
La variance : Un outil fondamental

- Comparer deux variables



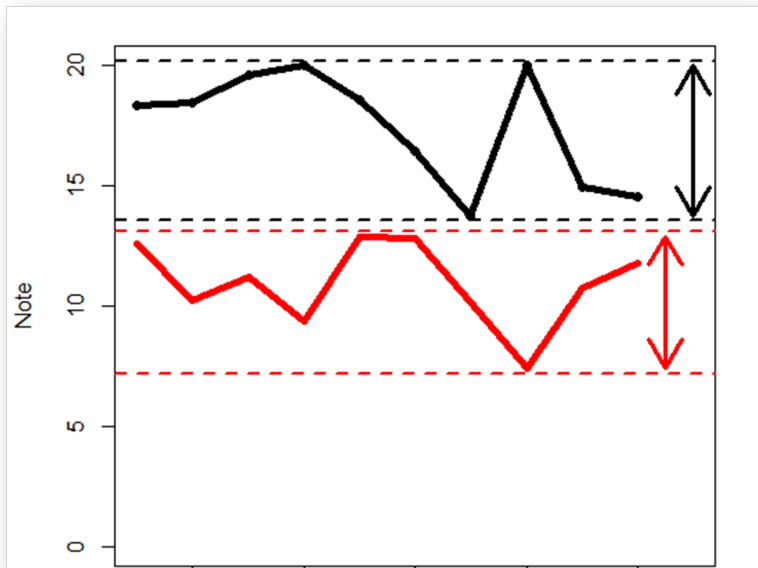
La variance : Un outil fondamental

► Comparer deux variables



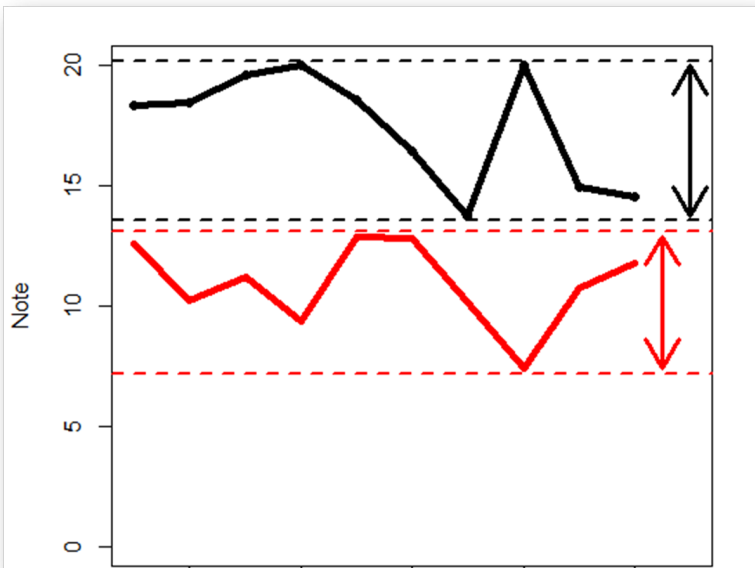
La variance : Un outil fondamental

- Comparer deux variables



La variance : Un outil fondamental

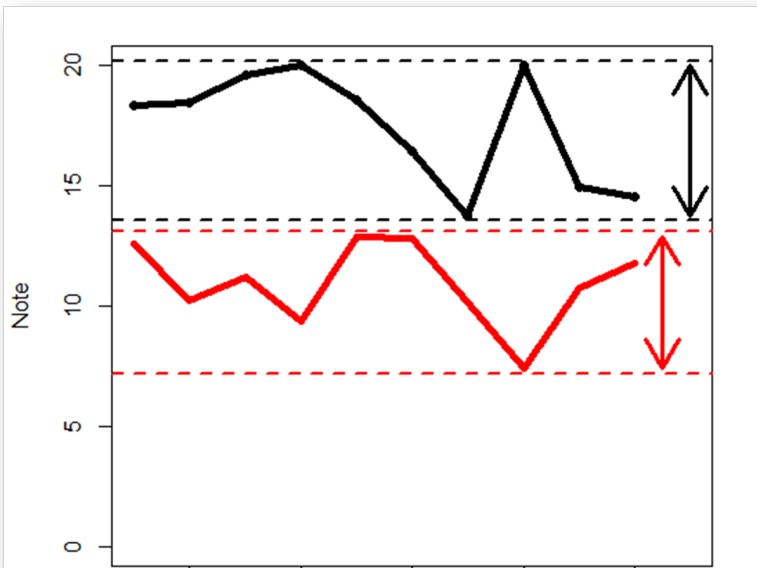
- Comparer deux variables



La variance : Un outil fondamental

- Comparer deux variables





La variance : Un outil fondamental

► Comparer deux variables

# Population et Échantillon

## La population complète

- Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE



## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

Échantillon

- ▶ **Le plus grand possible !!!**

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:



## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?



## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?

## La population complète

- ▶ Dans l'idéal, les comparaisons sont effectuées sur la population TOTALE
- ▶ **Problème:** Impossible de connaître TOUS les patients



Échantillon

## Échantillon

- ▶ **Le plus grand possible !!!**
- ▶ La précision des estimations augmente avec la taille d'échantillon.
- ▶ Exemple: Connaître le taux de vote pour un candidat aux présidentielles en interrogeant:
  - ▶ 4 personnes ?
  - ▶ 4590 personnes ?



# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.

## Variable qualitative

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)

## Variable qualitative

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON



# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme



# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme

# Variables quantitatives et qualitatives

## Variable quantitative

- ▶ Variable qu'il est possible de quantifier, de mesurer.
- ▶ Existe en deux formats: -continues: 12.4, 8.765, ( Taille, Pression artérielle etc...)
  - ▶ Discrètes : 2, 8, 14, 165 ( Nombre de lésions, rechutes, accouchements, etc..)

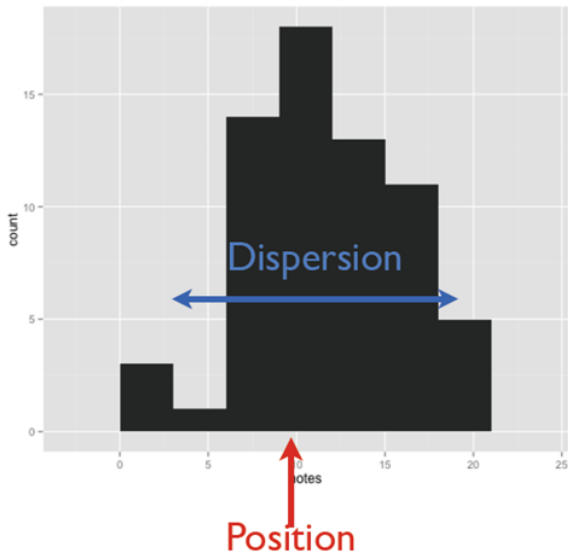
## Variable qualitative

- ▶ Variable **non dénombrable**
- ▶ Les niveaux de la variable sont appelés **modalités**
- ▶ Nominale / ordinales / binaires
- ▶ Exemples:
  - ▶ OUI / NON
  - ▶ Un peu / beaucoup / passionnément / ...
  - ▶ Bleu / Vert / Rouge
  - ▶ Homme / Femme

# Les statistiques descriptives

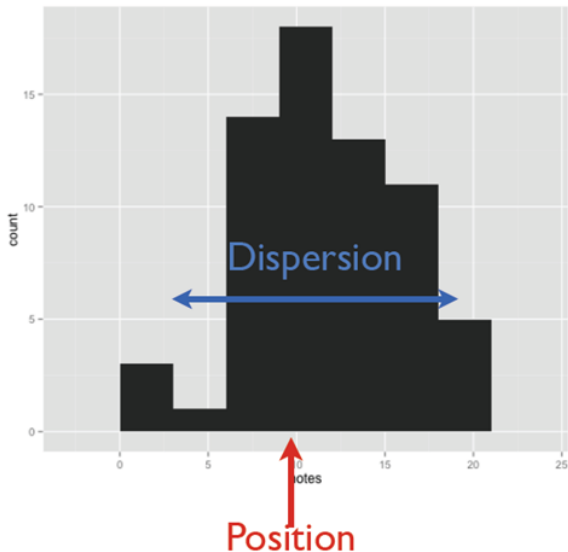
# Variables Quantitatives

## Position et Dispersion



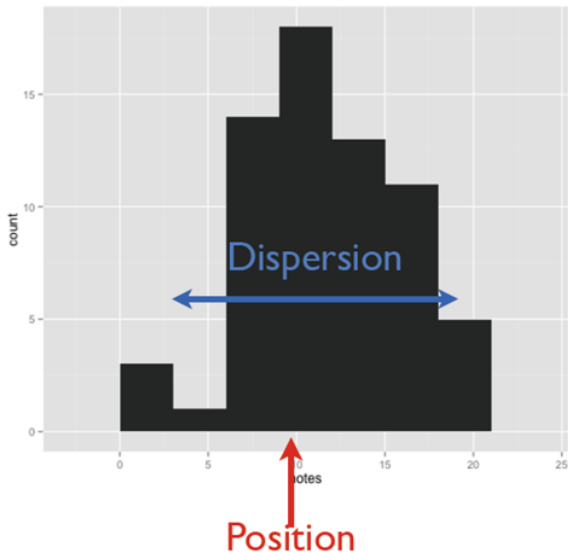
# Variables Quantitatives

## Position et Dispersion



# Variables Quantitatives

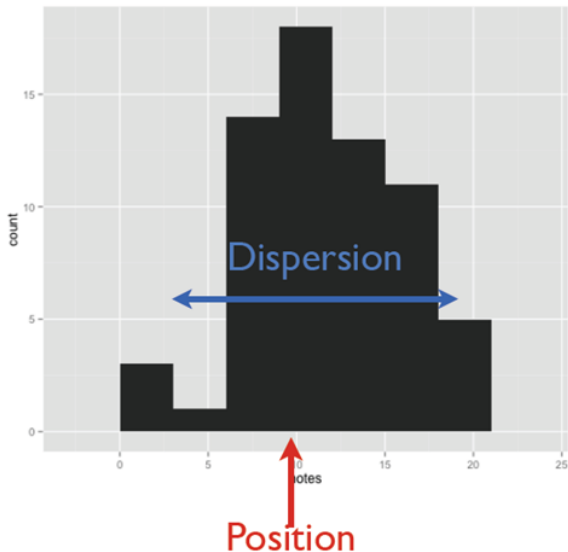
## Position et Dispersion





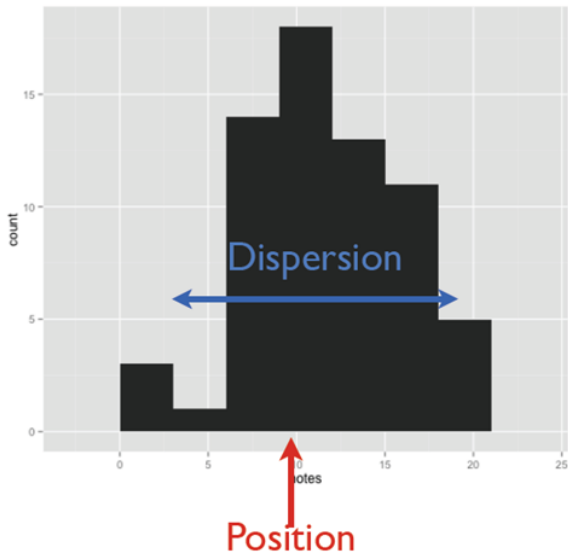
# Variables Quantitatives

## Position et Dispersion



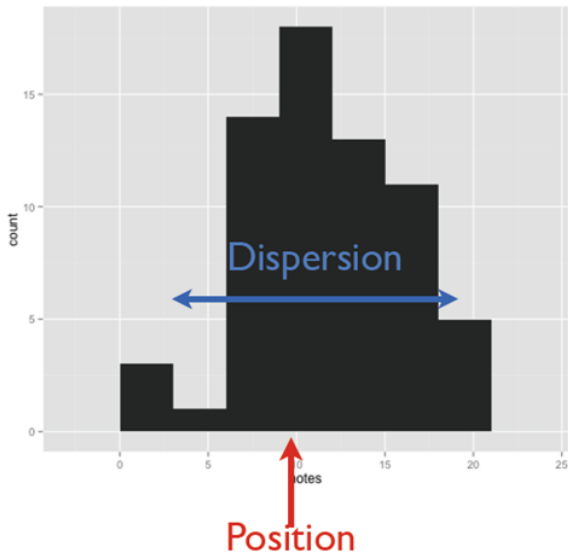
# Variables Quantitatives

## Position et Dispersion



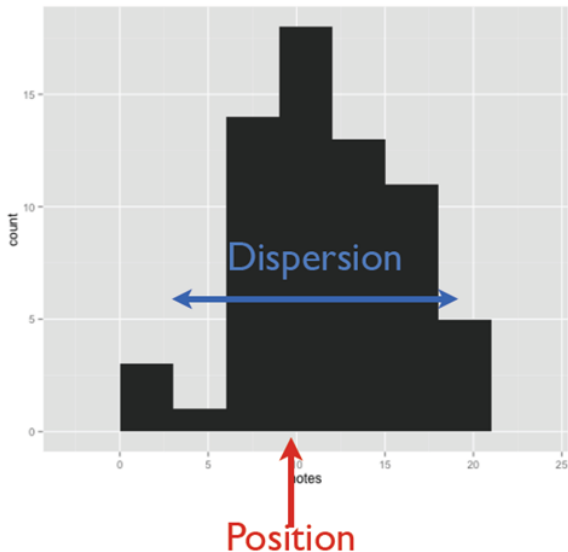
# Variables Quantitatives

## Position et Dispersion



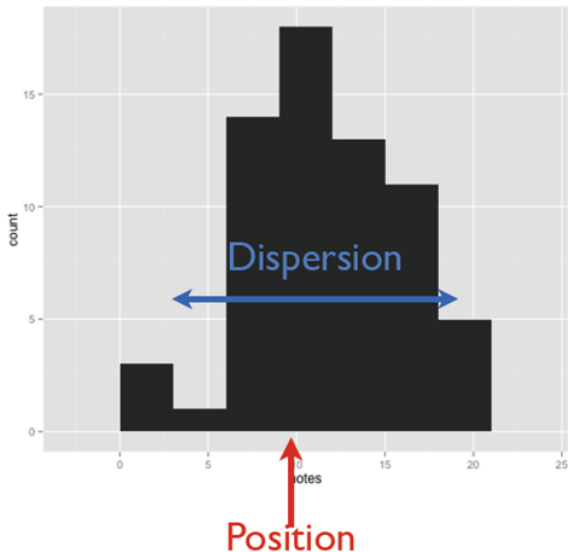
# Variables Quantitatives

## Position et Dispersion



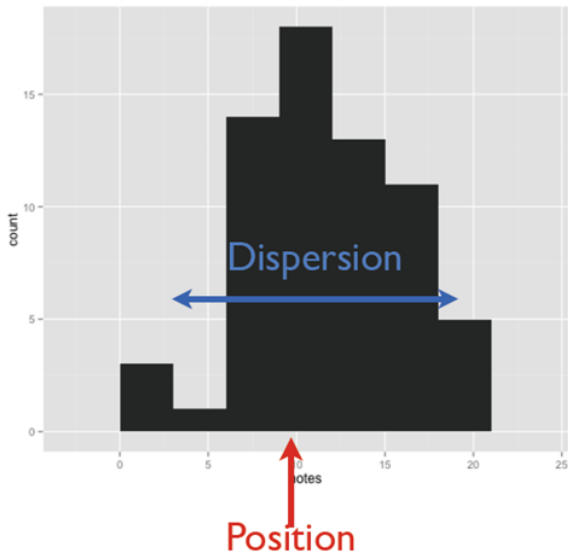
# Variables Quantitatives

## Position et Dispersion



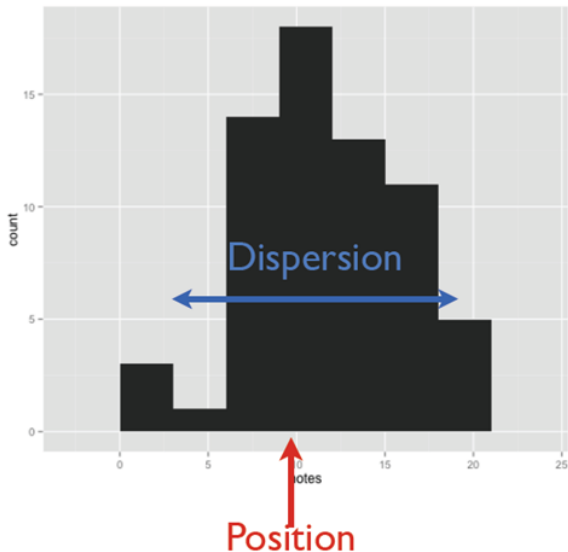
# Variables Quantitatives

## Position et Dispersion



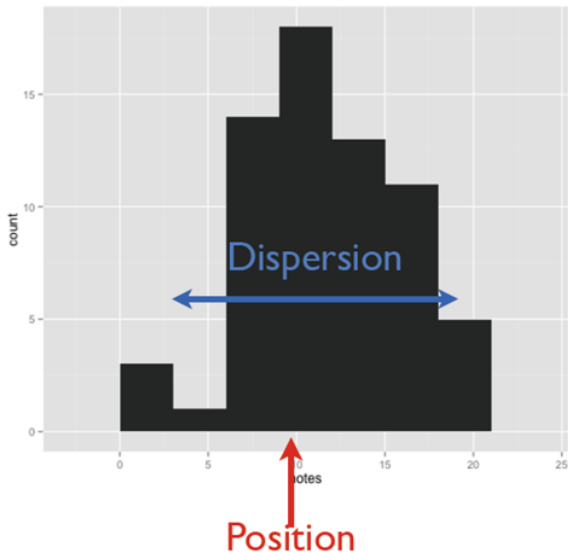
# Variables Quantitatives

## Position et Dispersion



# Variables Quantitatives

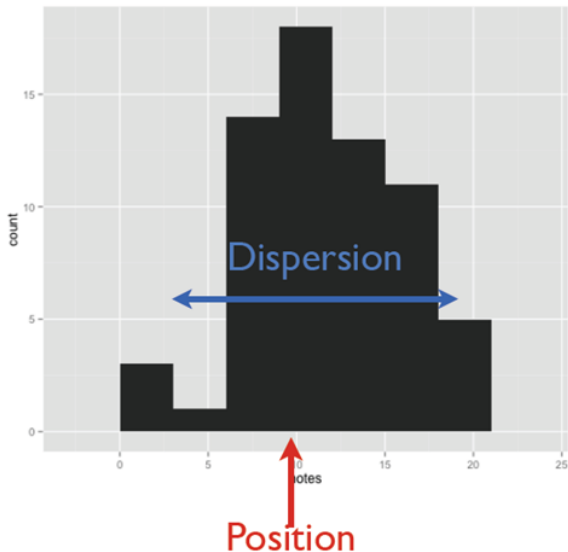
## Position et Dispersion





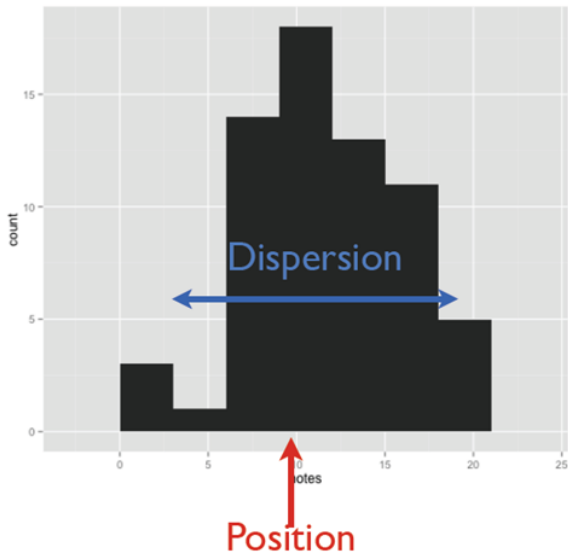
# Variables Quantitatives

## Position et Dispersion



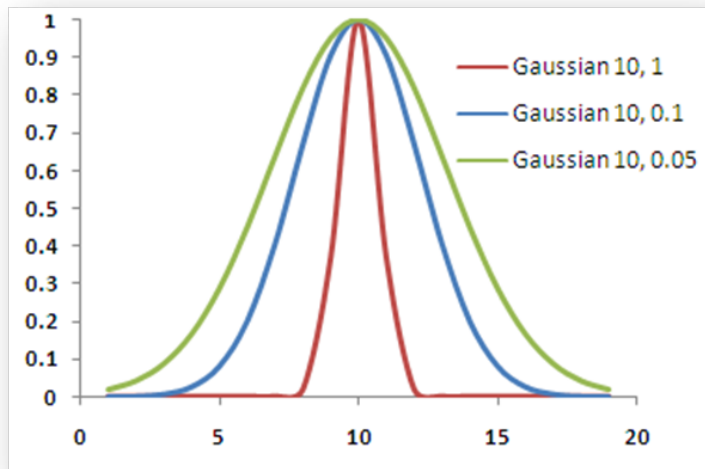
# Variables Quantitatives

## Position et Dispersion



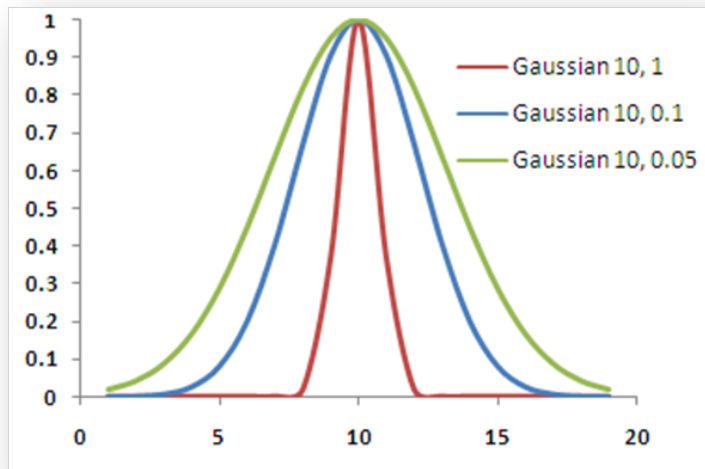
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



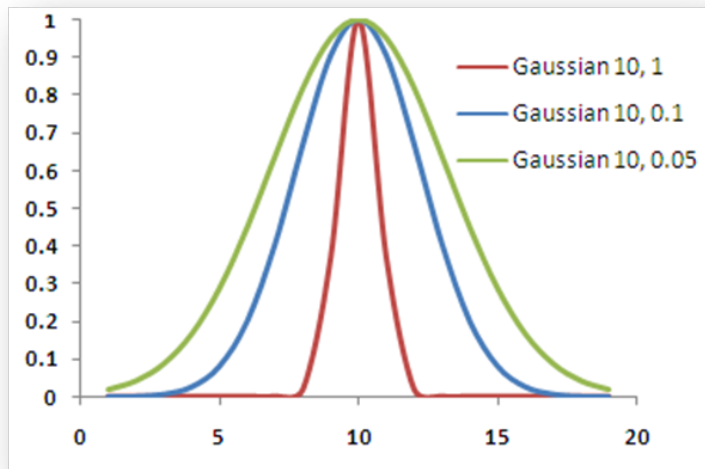
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



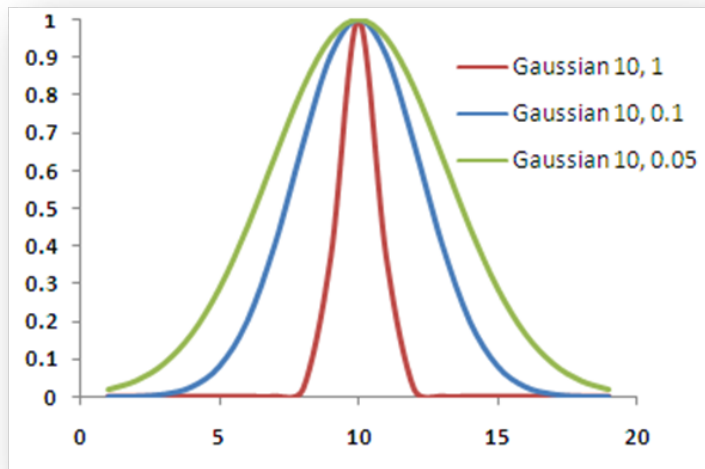
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



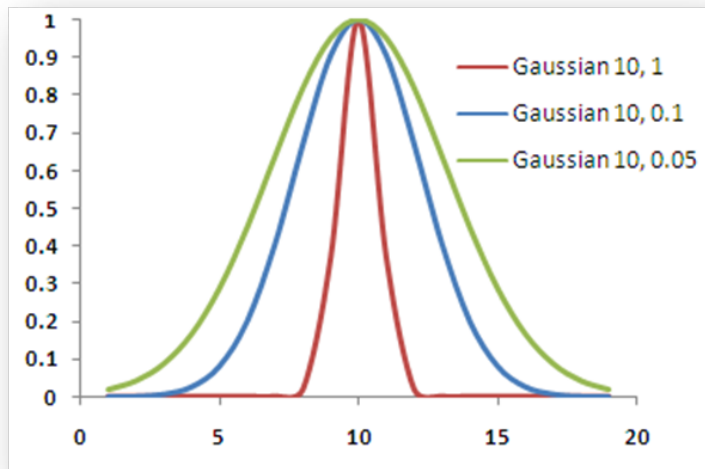
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



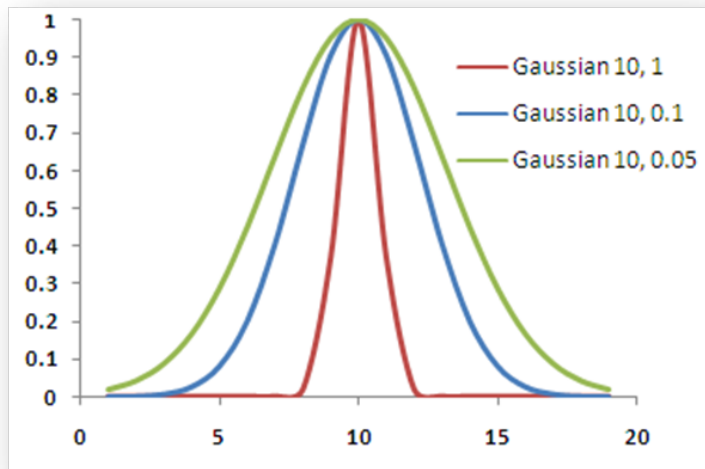
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



## Paramètre de dispersion

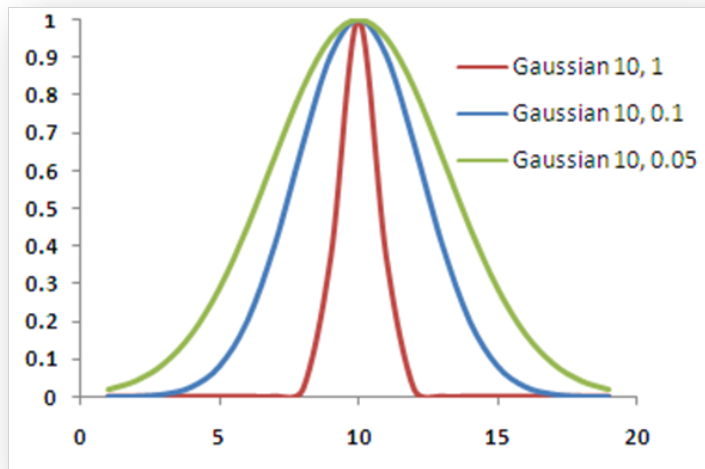
- Le paramètre de position ne suffit pas





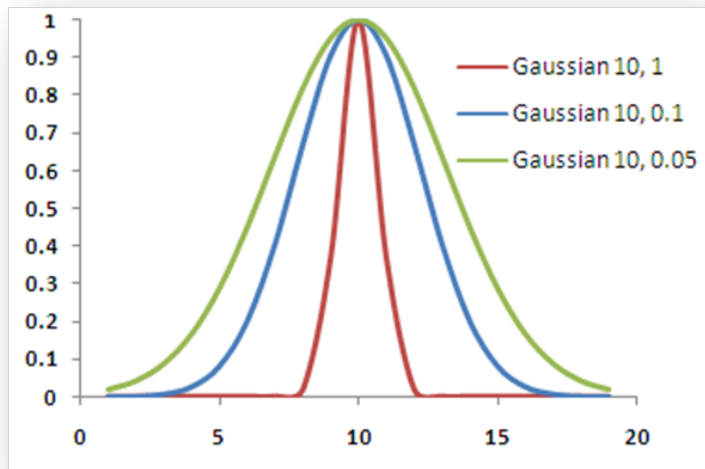
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



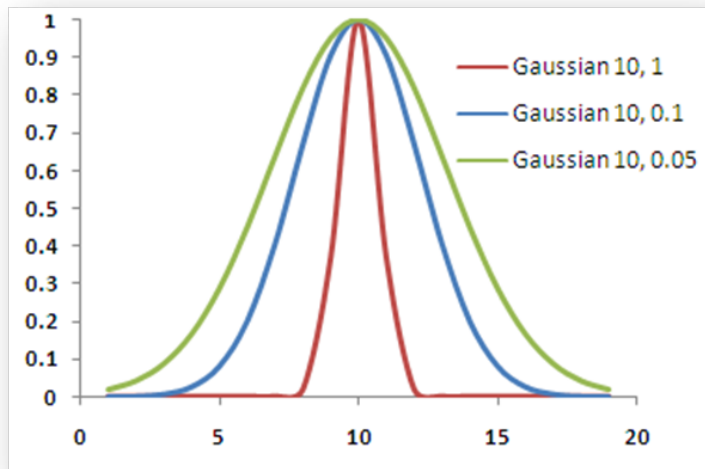
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



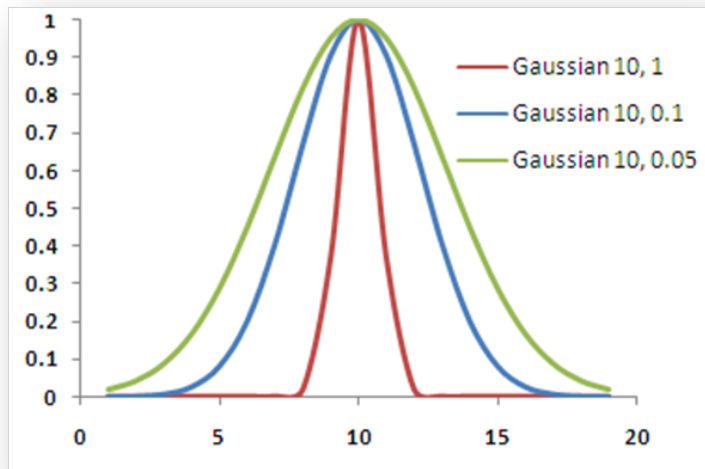
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



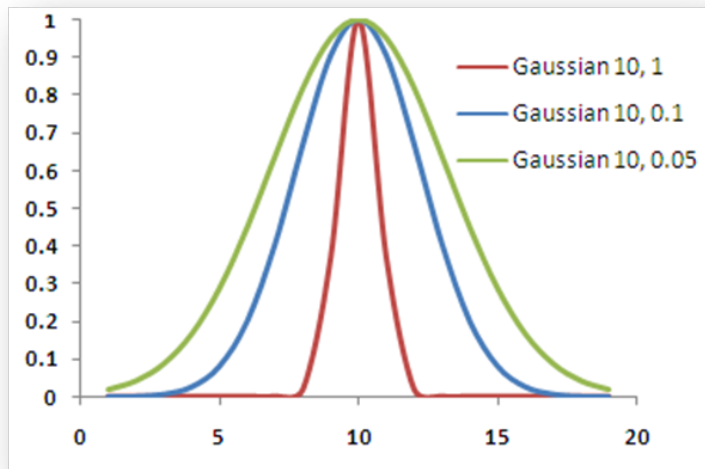
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



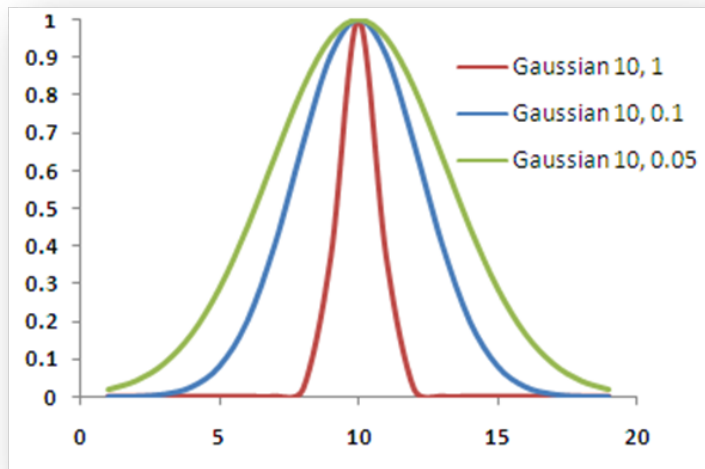
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



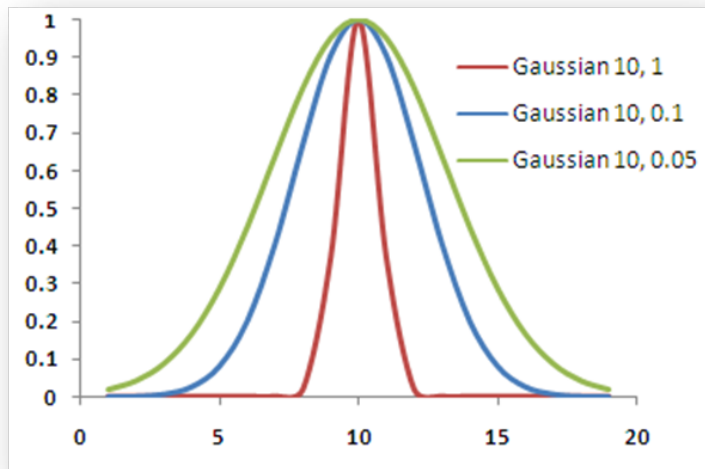
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



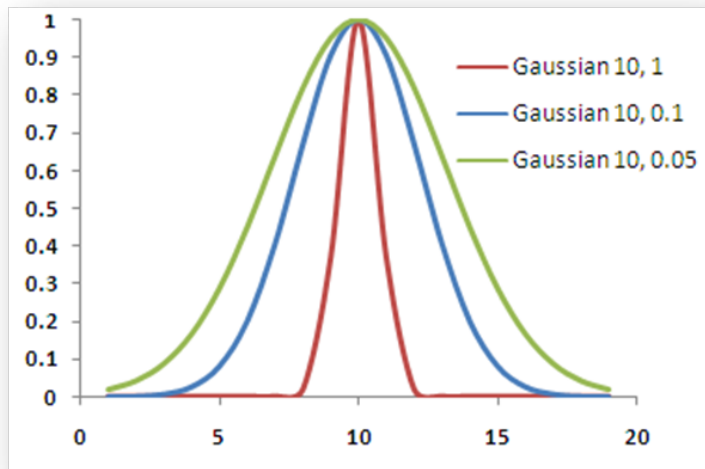
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



## Paramètre de dispersion

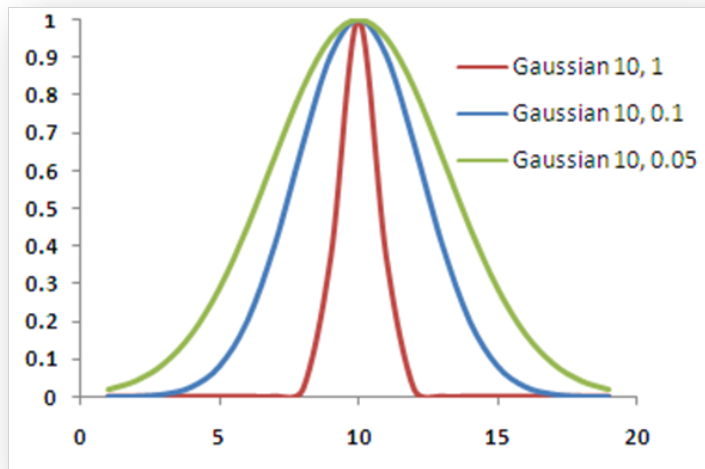
- Le paramètre de position ne suffit pas





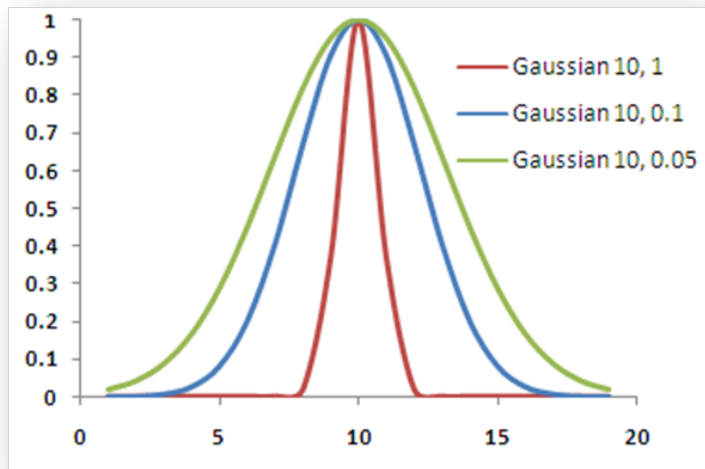
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



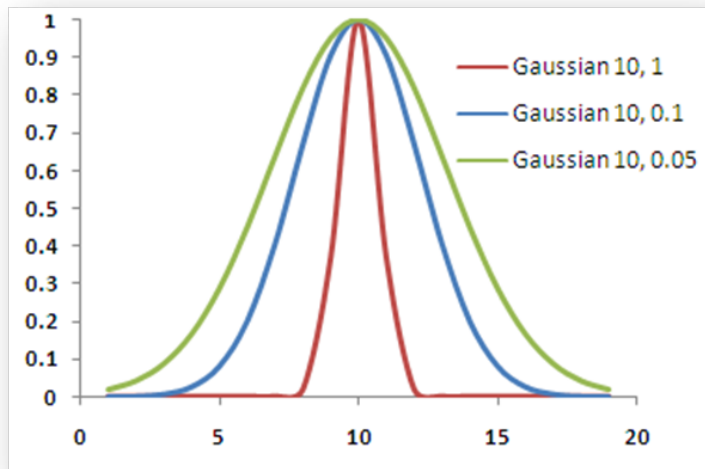
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



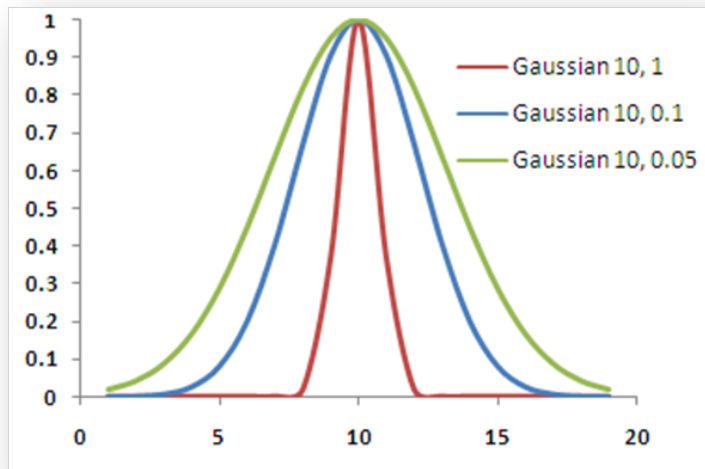
## Paramètre de dispersion

- Le paramètre de position ne suffit pas



## Paramètre de dispersion

- Le paramètre de position ne suffit pas



# Variables qualitatives

## Fréquences relatives

- Fréquence relative: pour résumer une distribution : calculer les fréquences relatives des sujets porteurs de chaque modalité. Ces proportions comprises entre 0 et 1,

# Variables qualitatives

## Fréquences relatives

- Fréquence relative: pour résumer une distribution : calculer les fréquences relatives des sujets porteurs de chaque modalité. Ces proportions comprises entre 0 et 1,

# Variables qualitatives

## Fréquences relatives

- Fréquence relative: pour résumer une distribution : calculer les fréquences relatives des sujets porteurs de chaque modalité. Ces proportions comprises entre 0 et 1,

Stade cancer	N	Pi (%)
1	88	54
2	29	17,8
3	34	20,9
4	12	7,4
Total	163	100

```
table(X)  
prop.table(table(X))
```

# Croisements

## Croisement entre les variables Quantitatives

- ▶ Évaluation graphique (cf. cours suivant)



# Croisements

## Croisement entre les variables Quantitatives

- ▶ Évaluation graphique (cf. cours suivant)
- ▶ Coefficient de corrélation linéaire de Bravais Pearson

# Croisements

## Croisement entre les variables Quantitatives

- ▶ Évaluation graphique (cf. cours suivant)
- ▶ Coefficient de corrélation linéaire de Bravais Pearson
- ▶ 
$$\rho = \frac{\sum[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{(x_i - \bar{x})^2 * (y_i - \bar{y})^2}}$$

# Croisements

## Croisement entre les variables Quantitatives

- ▶ Évaluation graphique (cf. cours suivant)
- ▶ Coefficient de corrélation linéaire de Bravais Pearson
- ▶ 
$$\rho = \frac{\sum[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{(x_i - \bar{x})^2 * (y_i - \bar{y})^2}}$$
- ▶  $-1 \leq \rho \leq 1$

# Croisements

## Croisement entre les variables Quantitatives

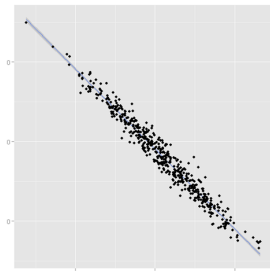
- ▶ Évaluation graphique (cf. cours suivant)
- ▶ Coefficient de corrélation linéaire de Bravais Pearson
- ▶ 
$$\rho = \frac{\sum[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{(x_i - \bar{x})^2 * (y_i - \bar{y})^2}}$$
- ▶  $-1 \leq \rho \leq 1$

# Croisements

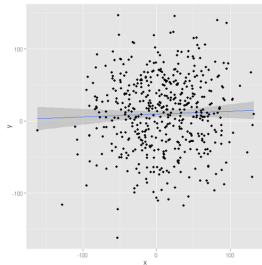
## Croisement entre les variables Quantitatives

- ▶ Évaluation graphique (cf. cours suivant)
- ▶ Coefficient de corrélation linéaire de Bravais Pearson
- ▶ 
$$\rho = \frac{\sum[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{(x_i - \bar{x})^2 * (y_i - \bar{y})^2}}$$
- ▶  $-1 \leq \rho \leq 1$

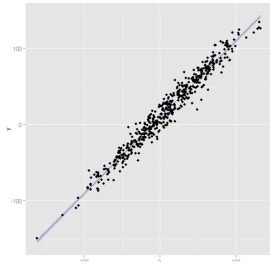
`cor(X,Y)`



$$\rho \approx -1$$



$$\rho \approx 0$$



$$\rho \approx 1$$



## Croisement entre les variables qualitatives

- ▶ Tableau de contingence



## Croisement entre les variables qualitatives

- ▶ Tableau de contingence

## Croisement entre les variables qualitatives

### ► Tableau de contingence

age / Sexe	H	F
0-5	4	4
5-10	2	8
1-18	4	11

## Croisement Quanti/Quali

### ► Mêmes éléments que description quanti mais par groupe

The data contains 150 observations, grouped by Species, of

- setosa (n = 50):

- Petal.Length: Mean = 1.46, SD = 0.17, range: [1, 1.90]
- Petal.Width: Mean = 0.25, SD = 0.11, range: [0.10, 0.60]

- versicolor (n = 50):