

# Statistiques et probabilités

## Cours n°3

Guillaume Postic

Université Paris-Saclay, Univ. Evry  
Département informatique

Master 1 MIAE - 2022/2023

# Variable aléatoire continue

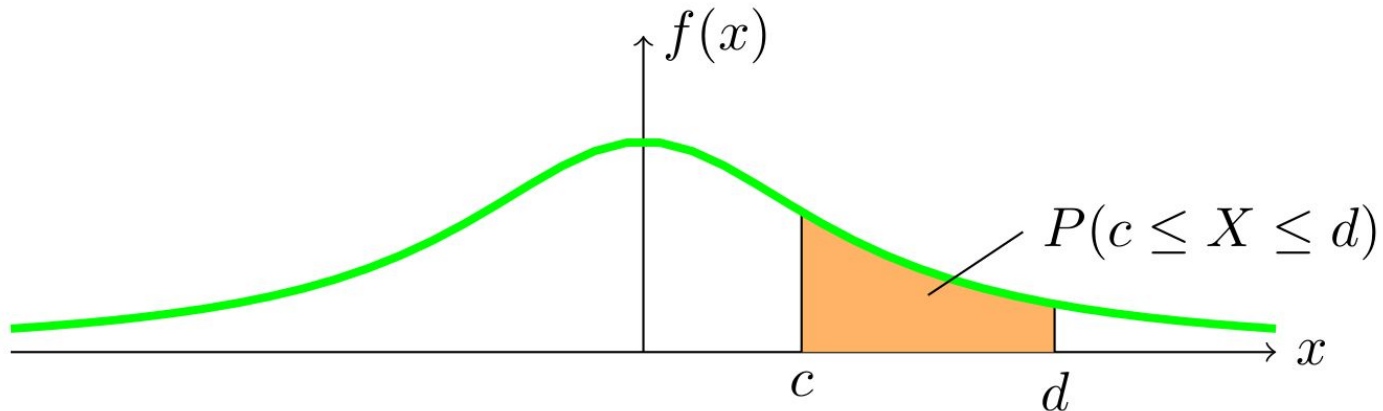
- Intervalle de valeurs continu :
  - e.g.  $[0, 1]$ ,  $[a, b]$ ,  $[0, \infty)$ ,  $(-\infty, \infty)$
- Fonction de densité ou densité de probabilité

$$f(x) \geq 0; \quad P(c \leq x \leq d) = \int_c^d f(x) dx$$

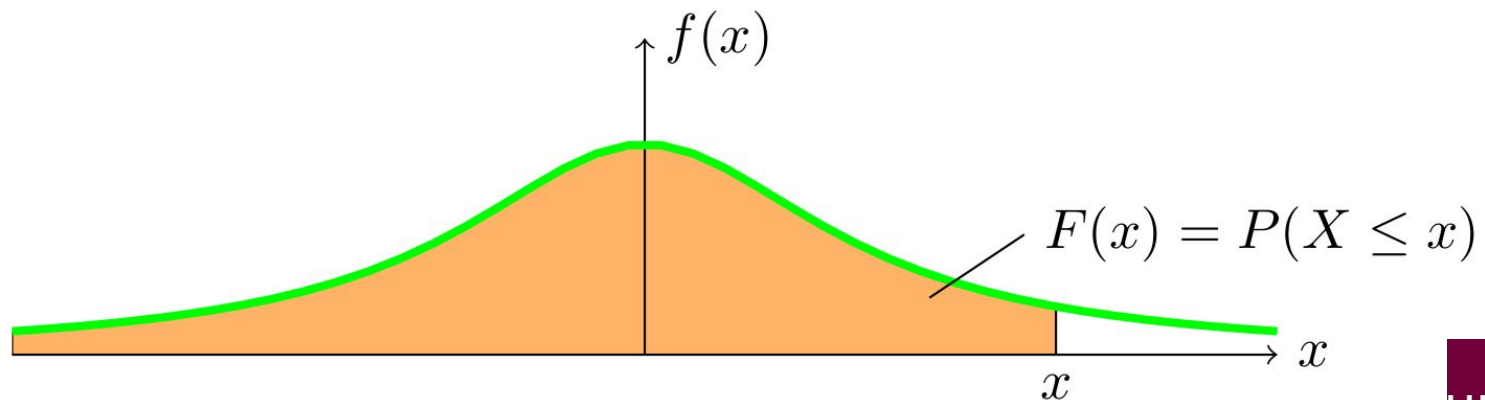
- Fonction de répartition

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$$

# Variable aléatoire continue



Fonction de densité et probabilité



Fonction de densité et fonction de répartition

# Propriétés de la fonction de répartition

(également valables pour les variables aléatoires discrètes)

- (Définition)  $F(x) = P(X \leq x)$
- $0 \leq F(x) \leq 1$
- Jamais décroissante
- $\lim_{x \rightarrow -\infty} F(x) = 0$
- $\lim_{x \rightarrow +\infty} F(x) = 1$
- $P(c < X \leq d) = F(d) - F(c)$
- $F'(x) = f(x)$

# Questions

Soit  $X$  une variable aléatoire continue.

- (a) Que vaut  $P(a \leq X \leq a)$  ?
- (b) Que vaut  $P(X = 0)$  ?
- (c) Est-ce que  $P(X = a) = 0$  signifie que  $X$  n'est jamais égal à  $a$  ?

# Espérance

$X$  continue sur  $[a, b]$ , avec la fonction de densité  $f(x)$  :

$$E(X) = \int_a^b xf(x) dx$$

$X$  discrète, de valeurs  $x_1, \dots, x_n$ , avec la fonction de masse  $p(x_i)$  :

$$E(X) = \sum_{i=1}^n x_i p(x_i)$$

# Variance et écart-type

Pour toute variable aléatoire  $X$  de moyenne  $\mu$

$$\text{Var}(X) = E((X - \mu)^2), \quad \sigma = \sqrt{\text{Var}(X)}$$

$X$  continue sur  $[a, b]$ , avec la fonction de densité  $f(x)$  :

$$\text{Var}(X) = \int_a^b (x - \mu)^2 f(x) dx.$$

$X$  discrète, de valeurs  $x_1, \dots, x_n$ , avec la fonction de masse  $p(x_i)$  :

$$\text{Var}(X) = \sum_{i=1}^n (x_i - \mu)^2 p(x_i).$$

# Quantiles (1)

Les quantiles sont les valeurs qui divisent un jeu de données en intervalles de probabilités égales.

Il y a donc un quantile de moins que le nombre de groupes créés. Par exemple, **les quartiles sont les trois quantiles** qui divisent un ensemble de données en quatre groupes de même probabilité.

La **médiane**, quant à elle, est le quantile qui sépare le jeu de données en deux groupes de même probabilité.



# Quantiles (2)

