Probabilités et Statistiques
I. Loi d'une variable continue

EPITA

Guillaume Euvrard

Variable aléatoire discrète ou continue

Imaginons un serveur qui reçoit des requêtes et considérons les variables aléatoires :

- $X =$ « Nombre de requêtes reçues en 1 minute ».
  X prend ses valeurs dans $\{0, 1, 2, \cdots, \}$.
- $T =$ « Temps d'attente avant la prochaine requête ».
  T prend ses valeurs dans $\mathbb{R}^+$.

On dit que X est une variable *discrète*, T une variable *continue*.

Variable aléatoire discrète ou continue

Étant donné un espace probabilisé (Ω, Σ, P), un variable aléatoire est une fonction $X : \Omega \longrightarrow \mathbb{R}$.

L'ensemble de ses valeurs possibles est $X(\Omega)$.

On s'intéresse à la **loi de** X : pour $A \subset X(\Omega)$, on veut définir $P(X \in A)$.

1. Quand on peut écrire $X(\Omega) = \{x_0, x_1, x_2, \cdots\}$, on dit que X est une variable discrète.
2. Si $X(\Omega)$ contient des intervalles, X est une variable continue.

Cas d'une variable discrète

Dans ce cas, $X(\Omega) = \{x_n, n \in K\}$ avec K de la forme $[\![1, N]\!]$ ou $K = \mathbb{N}$.

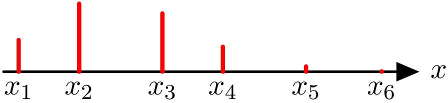
- Cas fini : $X(\Omega) = \{x_n, n \in [\![1, N]\!]\} = \{x_1, x_2, \cdots, x_N\}$
- Cas infini : $X(\Omega) = \{x_n, n \in \mathbb{N}\} = \{x_0, x_1, x_2, \cdots\}$

Cas d'une variable discrète

Dans ce cas, $X(\Omega) = \{x_n, n \in K\}$ avec K de la forme $[\![1, N]\!]$ ou $K = \mathbb{N}$.

- Cas fini : $X(\Omega) = \{x_n, n \in [\![1, N]\!]\} = \{x_1, x_2, \cdots, x_N\}$
- Cas infini : $X(\Omega) = \{x_n, n \in \mathbb{N}\} = \{x_0, x_1, x_2, \cdots\}$

On obtient la loi de X par la données des $P(X{=}x_n)$, $n \in K$.

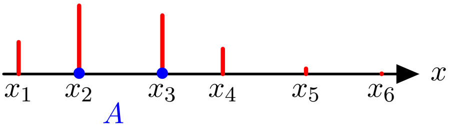


Cas d'une variable discrète

Dans ce cas, $X(\Omega) = \{x_n, n \in K\}$ avec K de la forme $[\![1, N]\!]$ ou $K = \mathbb{N}$.

- Cas fini : $X(\Omega) = \{x_n, n \in [\![1, N]\!]\} = \{x_1, x_2, \cdots, x_N\}$
- Cas infini : $X(\Omega) = \{x_n, n \in \mathbb{N}\} = \{x_0, x_1, x_2, \cdots\}$

On obtient la loi de X par la données des $P(X{=}x_n)$, $n \in K$.



Pour tout $A \subset X(\Omega)$

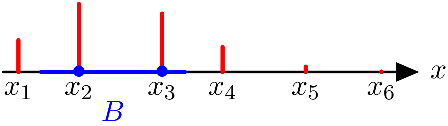
$$A = \{x_{k_1}, x_{k_2}, \cdots\} \qquad \text{et} \qquad P(X{\in}A) = \sum P(X{=}x_{k_n}\}$$

Cas d'une variable discrète

Dans ce cas, $X(\Omega) = \{x_n, n \in K\}$ avec K de la forme $[\![1, N]\!]$ ou $K = \mathbb{N}$.

- Cas fini : $X(\Omega) = \{x_n, n \in [\![1, N]\!]\} = \{x_1, x_2, \cdots, x_N\}$
- Cas infini : $X(\Omega) = \{x_n, n \in \mathbb{N}\} = \{x_0, x_1, x_2, \cdots\}$

On obtient la loi de X par la données des $P(X{=}x_n)$, $n \in K$.



Pour tout $A \subset X(\Omega)$ et pour $B \subset \mathbb{R}$, on prend $A = B \cap X(\Omega)$.

$$A = \{x_{k_1}, x_{k_2}, \cdots\} \qquad \text{et} \qquad P(X{\in}A) = \sum P(X{=}x_{k_n}\}$$

Cas d'une variable discrète

Dans ce cas, les évènements «$X \in \{x_n\}$» *engendrent* $\mathcal{P}\big(X(\Omega)\big)$. Ils permettent de définir

$$P : \left\{ \begin{array}{ccc} \mathcal{P}\big(X(\Omega)\big) & \longrightarrow & \mathbb{R} \\ A & \longmapsto & P(X{\in}A) \end{array} \right.$$

On en déduit $P(X{\in}A)$ pour tout $A \subset X(\Omega)$, puis pour tout $A \subset \mathbb{R}$.

Cas d'une variable discrète

Dans ce cas, les évènements «$X \in \{x_n\}$» *engendrent* $\mathcal{P}\big(X(\Omega)\big)$. Ils permettent de définir

$$P : \left\{ \begin{array}{ccc} \mathcal{P}\big(X(\Omega)\big) & \longrightarrow & \mathbb{R} \\ A & \longmapsto & P(X{\in}A) \end{array} \right.$$

On en déduit $P(X{\in}A)$ pour tout $A \subset X(\Omega)$, puis pour tout $A \subset \mathbb{R}$.

Il y a une contrainte : $\displaystyle\sum_{n \in K} P(X{=}x_n) = 1$

Cas d'une variable continue

Définition

Une variable aléatoire X est continue si l'ensemble $X(\Omega)$ de ses valeurs possibles est un intervalle de $\mathbb{R}$.

Cas d'une variable continue

Définition

Une variable aléatoire X est continue si l'ensemble $X(\Omega)$ de ses valeurs possibles est un intervalle de $\mathbb{R}$.

Théorème

Aucun intervalle de $\mathbb{R}$ n'est dénombrable.

On ne peut donc pas écrire $X(\Omega) = \{x_1, x_2, x_3, \cdots\}$.

Définition

Une variable aléatoire X est continue si l'ensemble $X(\Omega)$ de ses valeurs possibles est un intervalle de $\mathbb{R}$.

Théorème

Aucun intervalle de $\mathbb{R}$ n'est dénombrable.

On ne peut donc pas écrire $X(\Omega) = \{x_1, x_2, x_3, \cdots\}$.

Pour s'en sortir, on ne va pas chercher à définir $P(X \in A)$ pour tout $A \subset \mathbb{R}$.

Exemple : si on pose $P\big(X \in]-\infty, 3]\big) = 0.3$, on peut déduire

$$P(X \in \emptyset) = 0 \quad P\big(X \in]3, +\infty[\big) = 0.7 \quad P(X \in \mathbb{R}) = 1$$

On dit que $\Sigma = \big\{\emptyset,\ ]-\infty, 3],\ ]3, +\infty[, \mathbb{R}\big\}$ est une *tribu*.

Définition

Une tribu de $\mathbb{R}$ est un ensemble $\Sigma \subset \mathcal{P}(\mathbb{R})$ tel que :

1. $\mathbb{R} \in \Sigma$
2. Pour tout $A \in \Sigma$, $\overline{A} \in \Sigma$
3. Pour toute famille $(A_n, n \in K)$ de Σ (avec $K \subset \mathbb{N}$), $\bigcup_{n \in K} A_n \in \Sigma$

Définition

Une tribu de $\mathbb{R}$ est un ensemble $\Sigma \subset \mathcal{P}(\mathbb{R})$ tel que :

1. $\mathbb{R} \in \Sigma$
2. Pour tout $A \in \Sigma$, $\overline{A} \in \Sigma$
3. Pour toute famille $(A_n, n \in K)$ de Σ (avec $K \subset \mathbb{N}$), $\displaystyle\bigcup_{n \in K} A_n \in \Sigma$

Exemple : $\Sigma = \{\emptyset,]-\infty, 3],]3, +\infty[, \mathbb{R}\}$.

Σ est la plus petite tribu qui contient $]-\infty, 3]$. C'est la tribu *engendrée* par $]-\infty, 3]$.

Un peu de formalisme

Définition

Une tribu de $\mathbb{R}$ est un ensemble $\Sigma \subset \mathcal{P}(\mathbb{R})$ tel que :

1. $\mathbb{R} \in \Sigma$
2. Pour tout $A \in \Sigma$, $\overline{A} \in \Sigma$
3. Pour toute famille $(A_n, n \in K)$ de Σ (avec $K \subset \mathbb{N}$), $\displaystyle\bigcup_{n \in K} A_n \in \Sigma$

Définition

Une loi de probabilité sur Σ est une fonction $P : \Sigma \longrightarrow \mathbb{R}^+$ telle que :

1. $P(X \in \mathbb{R}) = 1$
2. Pour toute famille disjointe $(A_n, n \in K)$ de Σ (avec $K \subset \mathbb{N}$),
$$P\left(X \in \bigsqcup_{n \in K} A_n\right) = \sum_{n \in K} P(X \in A_n)$$

Exemples

Définition

Une loi de probabilité sur Σ est une fonction $P : \Sigma \longrightarrow \mathbb{R}^+$ telle que :

1. $P(X \in \mathbb{R}) = 1$
2. Pour toute famille disjointe $(A_n, n \in K)$ de Σ (avec $K \subset \mathbb{N}$),

$$P\left(X \in \bigsqcup_{n \in K} A_n\right) = \sum_{n \in K} P(X \in A_n)$$

- $\Sigma = \{\emptyset,]-\infty, 3],]3, +\infty[, \mathbb{R}\}$ et posons $P\big(X \in]-\infty, 3]\big) = 0.3$

  On déduit : $P(X \in \emptyset) = 0 \quad P\big(X \in]3, +\infty[\big) = 0.7 \quad P(X \in \mathbb{R}) = 1$

  On dit que Σ est la tribu *engendrée* par $]-\infty, 3]$.

Définition

Une loi de probabilité sur Σ est une fonction $P : \Sigma \longrightarrow \mathbb{R}^+$ telle que :

1. $P(X \in \mathbb{R}) = 1$
2. Pour toute famille disjointe $(A_n, n \in K)$ de Σ (avec $K \subset \mathbb{N}$),
$$P\left(X \in \bigsqcup_{n \in K} A_n\right) = \sum_{n \in K} P(X \in A_n)$$

- $\Sigma = \{\emptyset,]-\infty, 3],]3, +\infty[, \mathbb{R}\}$ et posons $P\big(X \in]-\infty, 3]\big) = 0.3$

  On déduit : $P(X \in \emptyset) = 0 \quad P\big(X \in]3, +\infty[\big) = 0.7 \quad P(X \in \mathbb{R}) = 1$

  On dit que Σ est la tribu *engendrée* par $]-\infty, 3]$.
- Tribu engendrée par $\{]-\infty, 3],]-\infty, 5]\}$?

  Que peut-on déduire de $P\big(X \in]-\infty, 3]\big)$ et $P\big(X \in]-\infty, 5]\big)$?

Définition

1. La tribu borélienne est la tribu Σ engendrée par tous les intervalles $]-\infty, x]$, $x \in \mathbb{R}$.
2. La fonction de répartition d'une variable aléatoire X est la fonction

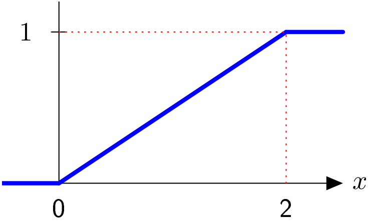
$$F_X : \begin{cases} \mathbb{R} & \longrightarrow & [0,1] \\ x & \longmapsto & P\big(X \in]-\infty, x]\big) \end{cases}$$

Définition

1. La tribu borélienne est la tribu Σ engendrée par tous les intervalles $]-\infty, x]$, $x \in \mathbb{R}$.
2. La fonction de répartition d'une variable aléatoire X est la fonction

$$F_X : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & [0,1] \\ x & \longmapsto & P\big(X \in]-\infty, x]\big) \end{array} \right.$$

Exemple : $F_X(x) = \left| \begin{array}{ll} 0 & \text{si } x < 0 \\ \frac{x}{2} & \text{si } 0 \leqslant x < 2 \\ 1 & \text{si } 2 \leqslant x \end{array} \right.$

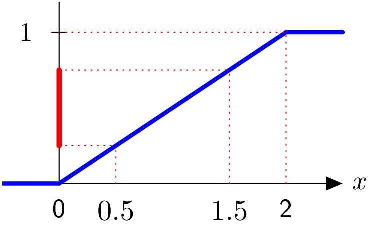


Définition

1. La tribu borélienne est la tribu Σ engendrée par tous les intervalles $]-\infty, x]$, $x \in \mathbb{R}$.
2. La fonction de répartition d'une variable aléatoire X est la fonction

$$F_X : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & [0,1] \\ x & \longmapsto & P(X \in]-\infty, x]) \end{array} \right.$$

Exemple : $F_X(x) = \left| \begin{array}{ll} 0 & \text{si } x < 0 \\ \frac{x}{2} & \text{si } 0 \leqslant x < 2 \\ 1 & \text{si } 2 \leqslant x \end{array} \right.$



$$P\left(X \in \left]\tfrac{1}{2}, \tfrac{3}{2}\right]\right) = F_X\left(\tfrac{3}{2}\right) - F_X\left(\tfrac{1}{2}\right) = \tfrac{1}{2}$$

Fonction de répartition : propriétés

La fonction de répartition F vérifie :

1. F est croissante
2. $\lim_{x \to -\infty} F(x) = 0$ et $\lim_{x \to +\infty} F(x) = 1$

Fonction de répartition : propriétés

La fonction de répartition F vérifie :

1. F est croissante
2. $\lim_{x \to -\infty} F(x) = 0$ et $\lim_{x \to +\infty} F(x) = 1$

Pour tout $(a, b) \in \mathbb{R}^2$ tel que $a \leqslant b$,

1. $]a, +\infty[\in \Sigma$ et $P\big(X \in]a, +\infty[\big) = 1 - F(a)$
2. $]a, b] \in \Sigma$ et $P\big(X \in]a, b]\big) = F(b) - F(a)$

Fonction de répartition : propriétés

La fonction de répartition F vérifie :

1. F est croissante
2. $\lim\limits_{x \to -\infty} F(x) = 0$ et $\lim\limits_{x \to +\infty} F(x) = 1$

Pour tout $(a, b) \in \mathbb{R}^2$ tel que $a \leqslant b$,

1. $]a, +\infty[\in \Sigma$ et $P\big(X \in]a, +\infty[\big) = 1 - F(a)$
2. $]a, b] \in \Sigma$ et $P\big(X \in]a, b]\big) = F(b) - F(a)$
3. $]-\infty, a[\in \Sigma$ et $P\big(X \in]-\infty, a[\big) = \lim\limits_{x \to a^-} F(x)$

Fonction de répartition : propriétés

La fonction de répartition F vérifie :

1. F est croissante
2. $\lim\limits_{x\to-\infty} F(x) = 0$ et $\lim\limits_{x\to+\infty} F(x) = 1$

Pour tout $(a,b) \in \mathbb{R}^2$ tel que $a \leqslant b$,

1. $]a,+\infty[\in \Sigma$ et $P\big(X\in]a,+\infty[\big) = 1 - F(a)$
2. $]a,b] \in \Sigma$ et $P\big(X\in]a,b]\big) = F(b) - F(a)$
3. $]-\infty,a[\in \Sigma$ et $P\big(X\in]-\infty,a[\big) = \lim\limits_{x\to a^-} F(x)$
4. $[a,+\infty[\in \Sigma$ et $P\big(X\in[a,+\infty[\big) = 1 - \lim\limits_{x\to a^-} F(x)$
5. $[a,b] \in \Sigma$ et $P\big(X\in[a,b]\big) = F(b) - \lim\limits_{x\to a^-} F(x)$

Fonction de répartition : propriétés

La fonction de répartition F vérifie :

1. F est croissante
2. $\lim_{x \to -\infty} F(x) = 0$ et $\lim_{x \to +\infty} F(x) = 1$

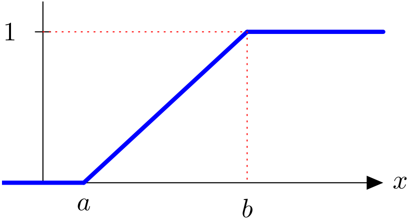
Pour tout $(a, b) \in \mathbb{R}^2$ tel que $a \leqslant b$,

1. $]a, +\infty[\in \Sigma$ et $P\big(X \in]a, +\infty[\big) = 1 - F(a)$
2. $]a, b] \in \Sigma$ et $P\big(X \in]a, b]\big) = F(b) - F(a)$
3. $]-\infty, a[\in \Sigma$ et $P\big(X \in]-\infty, a[\big) = \lim_{x \to a^-} F(x)$
4. $[a, +\infty[\in \Sigma$ et $P\big(X \in [a, +\infty[\big) = 1 - \lim_{x \to a^-} F(x)$
5. $[a, b] \in \Sigma$ et $P\big(X \in [a, b]\big) = F(b) - \lim_{x \to a^-} F(x)$

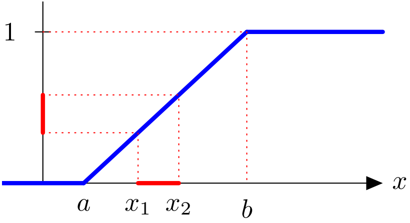
Remarque : si F continue en a, alors

$$P\big(X \in [a, a]\big) = F(a) - \lim_{x \to a^-} F(x) = 0$$

Fonction de répartition : exemples



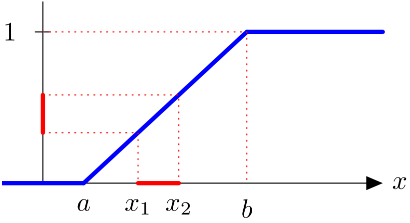
Si $x \in [a, b]$, $\quad F(x) = \dfrac{x-a}{b-a}$



Si $x \in [a, b]$, $\quad F(x) = \dfrac{x - a}{b - a}$

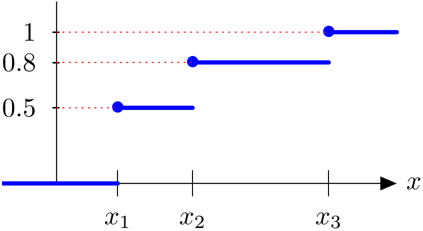
Ainsi, $P\big(X \in [x_1, x_2]\big) = \dfrac{x_2 - x_1}{b - a}$

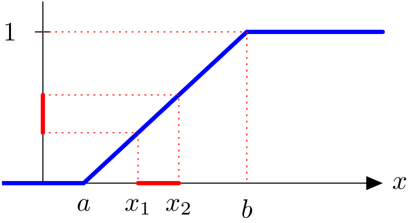
Fonction de répartition : exemples



Si $x \in [a, b]$, $\quad F(x) = \dfrac{x - a}{b - a}$

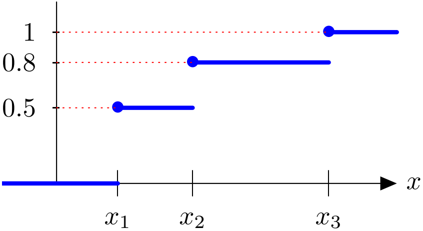
Ainsi, $P\big(X \in [x_1, x_2]\big) = \dfrac{x_2 - x_1}{b - a}$





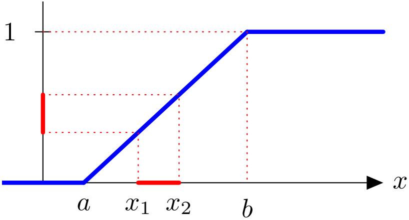
Si $x \in [a, b]$, $\quad F(x) = \dfrac{x-a}{b-a}$

Ainsi, $P\big(X \in [x_1, x_2]\big) = \dfrac{x_2 - x_1}{b-a}$



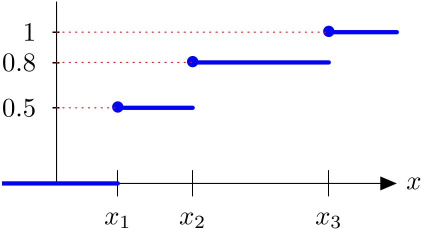
$$P\big(X{=}x_1\big) = F(x_1) - \lim_{x_1^-} F(x) = 0.5$$

Fonction de répartition : exemples



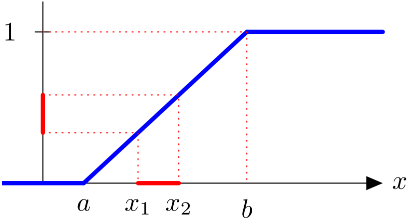
Si $x \in [a, b]$, $\quad F(x) = \dfrac{x-a}{b-a}$

Ainsi, $P\big(X \in [x_1, x_2]\big) = \dfrac{x_2 - x_1}{b-a}$



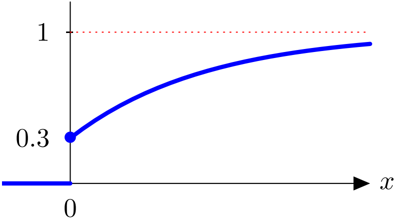
$$P\big(X{=}x_1\big) = F(x_1) - \lim_{x_1^-} F(x) = 0.5$$
$$P\big(X{=}x_2\big) = F(x_2) - \lim_{x_2^-} F(x) = 0.3$$
$$P\big(X{=}x_3\big) = F(x_3) - \lim_{x_3^-} F(x) = 0.2$$

Fonction de répartition : exemples



Si $x \in [a, b]$, $\quad F(x) = \dfrac{x-a}{b-a}$

Ainsi, $P\big(X \in [x_1, x_2]\big) = \dfrac{x_2 - x_1}{b-a}$



Fonction de densité

Soit X une variable aléatoire et F sa fonction de répartition.

Définition

S'il existe une fonction f telle que $F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$

alors f est une densité de X.

Soit X une variable aléatoire et F sa fonction de répartition.

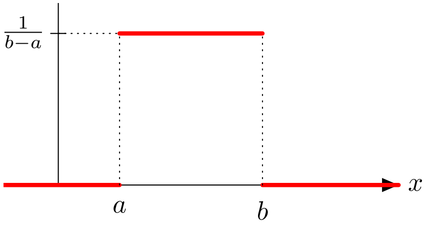
Définition

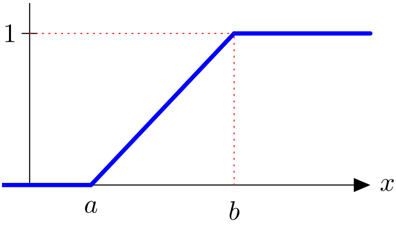
S'il existe une fonction f telle que $F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$

alors f est une densité de X.

Remarque : on a alors souvent $F'(x) = f(x)$. Mais *pas toujours*.

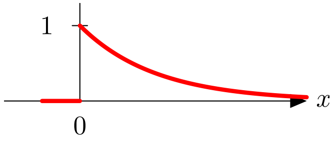
Exemple : loi uniforme sur $[a, b]$.



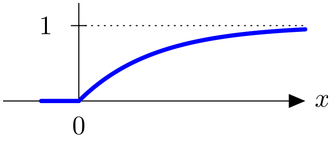


Exemple de densité : loi exponentielle

Densité f



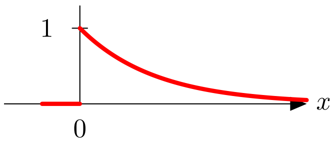
Répartition F



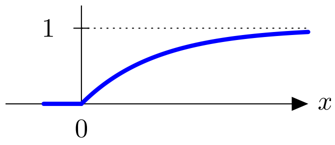
$$f(x) = \left| \begin{array}{cl} 0 & \text{si } x < 0 \\ e^{-x} & \text{si } x > 0 \end{array} \right.$$

Exemple de densité : loi exponentielle

Densité f



Répartition F

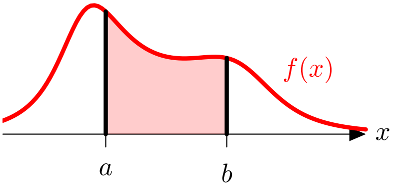


$$f(x) = \left| \begin{array}{ll} 0 & \text{si } x < 0 \\ e^{-x} & \text{si } x > 0 \end{array} \right.$$

$$F(x) = \left| \begin{array}{ll} 0 & \text{si } x < 0 \\ 1 - e^{-x} & \text{si } x \geqslant 0 \end{array} \right.$$

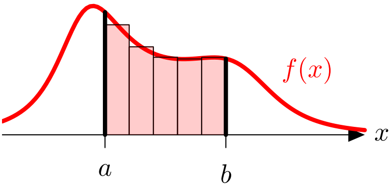
Interprétation

Vision exacte

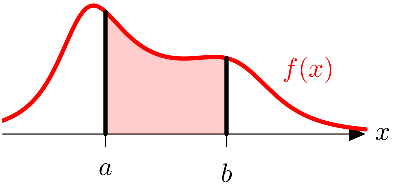


$$P\big(X \in [a,b]\big) = \int_a^b f(x)\,\mathsf{d}x$$

Vision approchée

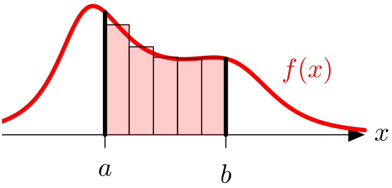


Vision exacte



$$P\big(X \in [a,b]\big) = \int_a^b f(x)\,\mathsf{d}x$$

Vision approchée



$$P\big(X \in [a,b]\big) \approx \sum_{x_k \in [a,b]} f(x_k)\delta x$$

Propriété de la densité

Soit X une variable aléatoire admettant une densité f.

1. f est positive.
2. L'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t$ converge et vaut 1.
3. Pour tout $[a,b] \subset \mathbb{R}$, $P(X \in [a,b]) = \displaystyle\int_a^b f(t)\,\mathrm{d}t$.

Propriété de la densité

Soit X une variable aléatoire admettant une densité f.

1. f est positive.
2. L'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t$ converge et vaut 1.
3. Pour tout $[a,b] \subset \mathbb{R}$, $\displaystyle P(X \in [a,b]) = \int_a^b f(t)\,\mathrm{d}t$.

Inversement, tout fonction f satisfaisant ces propriétés est une densité.

Propriété de la densité

Soit X une variable aléatoire admettant une densité f.

1. f est positive.
2. L'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t$ converge et vaut 1.
3. Pour tout $[a,b] \subset \mathbb{R}$, $P(X \in [a,b]) = \displaystyle\int_a^b f(t)\,\mathrm{d}t$.

Inversement, tout fonction f satisfaisant ces propriétés est une densité.

De plus la fonction de répartition F est continue. Ainsi, pour tout $[a,b] \subset \mathbb{R}$:

1. $P(X{=}a) = 0$
2. $P\big(X \in]a,b]\big) = P\big(X \in [a,b]\big) = \displaystyle\int_a^b f(t)\,\mathrm{d}t$

En conclusion

1. Pour décrire la distribution d'une variable continue X, on cherche les $P(X \in A)$, mais pas pour tout $A \subset \mathbb{R}$.

En conclusion

1. Pour décrire la distribution d'une variable continue X, on cherche les $P(X \in A)$, mais pas pour tout $A \subset \mathbb{R}$.
2. Toute variable aléatoire X admet une fonction de répartition : $F(x) = P\big(X \in]-\infty, x]\big) = P(X \leqslant x)$

En conclusion

1. Pour décrire la distribution d'une variable continue X, on cherche les $P(X \in A)$, mais pas pour tout $A \subset \mathbb{R}$.
2. Toute variable aléatoire X admet une fonction de répartition : $F(x) = P\big(X \in]-\infty, x]\big) = P(X \leqslant x)$
3. De nombreuses variables continues admettent une densité :

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

En conclusion

1. Pour décrire la distribution d'une variable continue X, on cherche les $P(X \in A)$, mais pas pour tout $A \subset \mathbb{R}$.
2. Toute variable aléatoire X admet une fonction de répartition : $F(x) = P\big(X \in]-\infty, x]\big) = P(X \leqslant x)$
3. De nombreuses variables continues admettent une densité :
$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$$
4. Pour tout $[a, b] \subset \mathbb{R}$, $P\big(X \in [a, b]\big) = \displaystyle\int_a^b f(t)\,\mathrm{d}t$