Processus stochastiques et mouvement Brownien

Une introduction informelle

Pierre Gloaguen

30/03/2020

Objectifs du cours

- Initiation à la modélisation stochastique de séries temporelles;
- Description de phénomènes dynamiques à l'aide des probabilités;

Objectifs du cours

- Initiation à la modélisation stochastique de séries temporelles;
- Description de phénomènes dynamiques à l'aide des probabilités;
- Nécessité de décrire un aléa temporel:
  - **Processus stochastiques**.

Objectifs du cours

- Initiation à la modélisation stochastique de séries temporelles;
- Description de phénomènes dynamiques à l'aide des probabilités;
- Nécessité de décrire un aléa temporel:
  - **Processus stochastiques**.
- Focalisation sur les modèles de dynamiques en temps continu.

Modèles d'équations différentielles

On s'intéresse à une quantité $x(t)$ définie continument au cours du temps:

- **Physique:** la position d'une particule
- **Ecologie:** la taille d'une population, déplacement d'un individu
- **Epidemiologie:** la proportion d'une population infect
- **Finance:** la valeur d'une action

Modèles d'équations différentielles

On s'intéresse à une quantité $x(t)$ définie continument au cours du temps:

- **Physique:** la position d'une particule
- **Ecologie:** la taille d'une population, déplacement d'un individu
- **Epidemiologie:** la proportion d'une population infect
- **Finance:** la valeur d'une action

Un modèle d'EDO est donné par une fonction f telle que $x(t)$ satisfait:

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = f(x(t), t), \; x(0) = x_0$$

La solution est une **fonction déterministe**:

$$x(t) = x(0) + \int_0^t f(x(s), s)\mathrm{d}s$$

Inclure la variabilité

Il n'est pas réaliste de vouloir penser certains phénomènes de manière déterministe:, on voudrait inclure de l'aléa (ou bruit):

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = f(x(t), t, \text{Aléa}),\ x(0) \sim \text{Aléa}$$

- Quelle structure d'aléa utiliser pour décrire les séries temporelles continues?
  - Un processus particulier: **Le mouvement Brownien**.
  - Cela aboutit aux modèles d'**équations différentielles stochastiques**.

Inclure la variabilité

Il n'est pas réaliste de vouloir penser certains phénomènes de manière déterministe:, on voudrait inclure de l'aléa (ou bruit):

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = f(x(t), t, \text{Aléa}),\ x(0) \sim \text{Aléa}$$

- Quelle structure d'aléa utiliser pour décrire les séries temporelles continues?
  - Un processus particulier: **Le mouvement Brownien**.
  - Cela aboutit aux modèles d'**équations différentielles stochastiques**.

Objectif

Découverte du mouvement Brownien et des équations différentielles stochastiques, et comment les simuler (avec le logiciel `R`).

Processus stochastiques

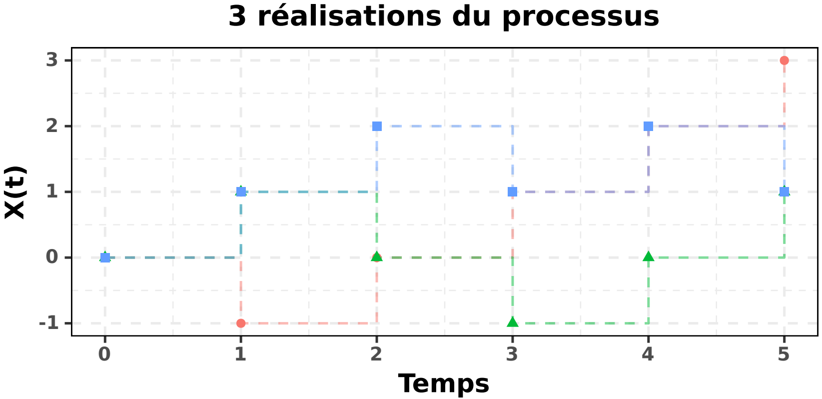
Définition

- Une collection de variables aléatoires indexées par le temps:
  - Temps discret, aux temps $0, 1, \ldots, n$, les V.A. $X(1), X(2), \ldots, X(n)$.
  - Temps continu, $t \in \mathbb{R}_+$, la trajectoire est noté $\{X(t)\}_{t \geq 0}$.
- $X(t)$ prend valeurs dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ ou $\mathbb{R}^d$.

Exemple en temps discret (marche aléatoire):

$$X(t+1) = X(t) + \zeta_t,\ \ X(0) = 0,\ \ t = 0, 1, \ldots$$

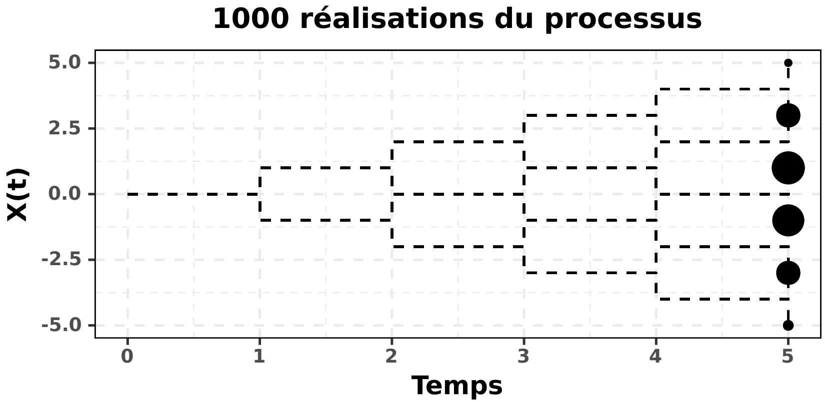
$$\zeta_t \overset{\text{ind.}}{=} \begin{cases} 1 & \text{avec proba. } \frac{1}{2} \\ -1 & \text{avec proba. } \frac{1}{2} \end{cases}$$



Exemple en temps discret (marche aléatoire):

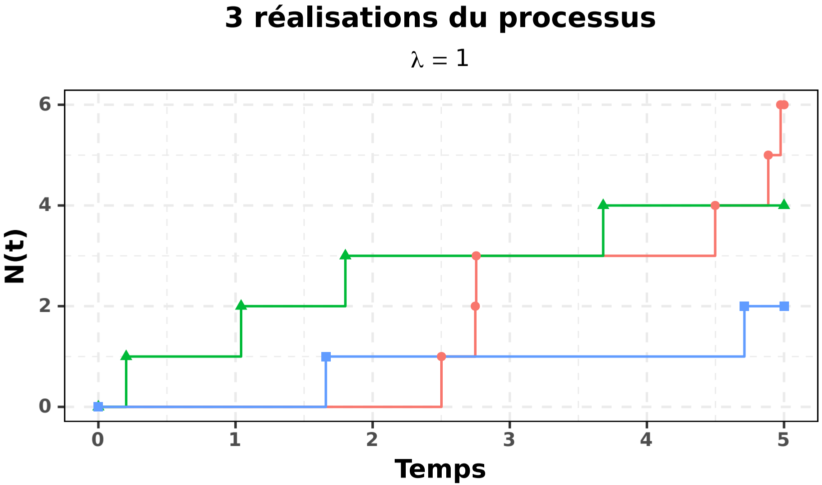
$$X(t+1) = X(t) + \zeta(t),\ \ X(0) = 0,\ \ t = 0, 1, \ldots$$

$$\zeta(t) \overset{\text{ind.}}{=} \begin{cases} 1 & \text{avec proba. } \frac{1}{2} \\ -1 & \text{avec proba. } \frac{1}{2} \end{cases}$$



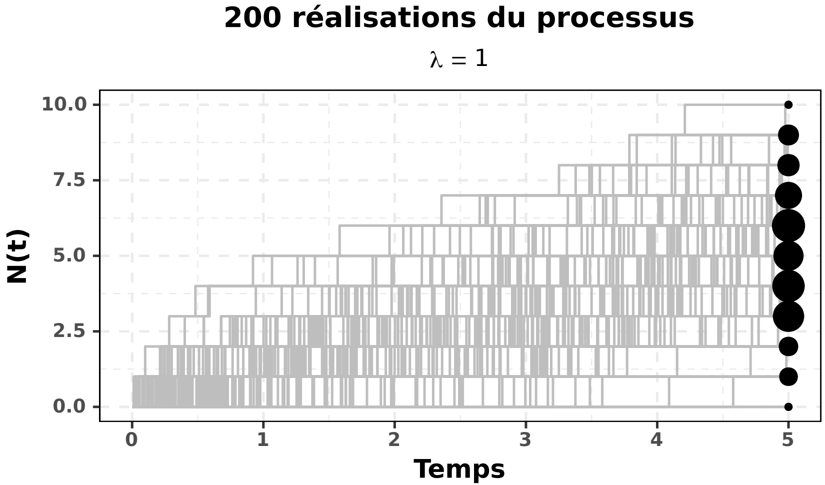
Ex. en temps continu (Processus de poisson):

- Instants de sauts $T_0 = 0, T_1, T_2, \ldots$
- Loi du temps entre deux sauts: $T_k - T_{k-1} \sim \mathcal{E}xp(\lambda)$
- Nombre de sauts jusqu'à l'instant t:
  $N(0) = 0,\ N(t) = \sum_{k=1}^{\infty} \mathbf{1}_{T_k \leq t}$



Ex. en temps continu (Processus de poisson):

- Instants de sauts $T_0 = 0, T_1, T_2, \ldots$
- Loi du temps entre deux sauts: $T_k - T_{k-1} \sim \mathcal{E}xp(\lambda)$
- Nombre de sauts jusqu'à l'instant t:
  $N(0) = 0,\ N(t) = \sum_{k=1}^{\infty} \mathbf{1}_{T_k \leq t}$



Processus Markovien

L'influence du passé se résume au dernier instant.

Un processus $\{X(t)\}_{t\geq 0}$ est Markovien si, pour toute suite de temps $t_1 < t_2 < \dot{<} t_n$, la loi de $X_{t_{n+1}}|X_{t_1},\ldots,X_{t_n}$ est égale à celle de $X_{t_{n+1}}|X_{t_n}$.

Propriété: Un processus Markovien est **entièrement caractérisé** par:

- La loi de $X(0)$: *la loi initiale*
- La loi de $X_{t_{n+1}}|X_{t_n}$: *la loi de transition*.

Processus Markovien

L'influence du passé se résume au dernier instant.

Un processus $\{X(t)\}_{t\geq 0}$ est Markovien si, pour toute suite de temps $t_1 < t_2 < \dots < t_n$, la loi de $X_{t_{n+1}}|X_{t_1},\dots,X_{t_n}$ est égale à celle de $X_{t_{n+1}}|X_{t_n}$.

Propriété: Un processus Markovien est **entièrement caractérisé** par:

- La loi de $X(0)$: *la loi initiale*
- La loi de $X_{t_{n+1}}|X_{t_n}$: *la loi de transition*.

Si on connaît ces deux lois, on peut simuler le processus aux temps $t_0 = 0, t_1, \dots, t_n$

- On simule $X(0)$ selon la loi initiale.
- On simule selon $X(t_1)|X(0)$, puis $X(t_2)|X(t_1)$, etc...

Loi de transition

- **Marche aléatoire**

$$X_{t+1}|X_t = \begin{cases} X_t + 1 & \text{avec proba } \frac{1}{2} \\ X_t - 1 & \text{avec proba } \frac{1}{2} \end{cases}$$

- **Processus de poisson** de paramètre λ:

$$X_{t+\Delta}|X_t = X_t + \mathcal{P}oisson(\lambda \times \Delta)$$

Mouvement Brownien

Définition

Un mouvement Brownien, noté $\{B(s)\}_{s\geq 0}$, est un processus stochastique défini en *temps continu* satisfaisant les hypothèses suivantes:

1. $B(0) = 0$.

Définition

Un mouvement Brownien, noté $\{B(s)\}_{s\geq 0}$, est un processus stochastique défini en *temps continu* satisfaisant les hypothèses suivantes:

1. $B(0) = 0$.
2. Les accroissements sont indépendants, c'est à dire, pour tous $0 \leq t_1 < t_2 < t_3$, les variables aléatoires $B(t_2) - B(t_1)$ et $B(t_3) - B_(t_2)$ sont indépendantes.

Définition

Un mouvement Brownien, noté $\{B(s)\}_{s\geq 0}$, est un processus stochastique défini en *temps continu* satisfaisant les hypothèses suivantes:

1. $B(0) = 0$.
2. Les accroissements sont indépendants, c'est à dire, pour tous $0 \leq t_1 < t_2 < t_3$, les variables aléatoires $B(t_2) - B(t_1)$ et $B(t_3) - B_(t_2)$ sont indépendantes.
3. Les accroissements sont stationnaires, c'est à dire que, pour tout $t > 0$ et $h > 0$, la loi de la variable aléatoire $B(t + h) - B(t)$ ne dépend que de h.

Définition

Un mouvement Brownien, noté $\{B(s)\}_{s\geq 0}$, est un processus stochastique défini en *temps continu* satisfaisant les hypothèses suivantes:

1. $B(0) = 0$.
2. Les accroissements sont indépendants, c'est à dire, pour tous $0 \leq t_1 < t_2 < t_3$, les variables aléatoires $B(t_2) - B(t_1)$ et $B(t_3) - B_(t_2)$ sont indépendantes.
3. Les accroissements sont stationnaires, c'est à dire que, pour tout $t > 0$ et $h > 0$, la loi de la variable aléatoire $B(t+h) - B(t)$ ne dépend que de h.
4. Pour tout $t > 0$, $B(t) \sim \mathcal{N}(0, t)$

Définition

Un mouvement Brownien, noté $\{B(s)\}_{s\geq 0}$, est un processus stochastique défini en *temps continu* satisfaisant les hypothèses suivantes:

1. $B(0) = 0$.
2. Les accroissements sont indépendants, c'est à dire, pour tous $0 \leq t_1 < t_2 < t_3$, les variables aléatoires $B(t_2) - B(t_1)$ et $B(t_3) - B_(t_2)$ sont indépendantes.
3. Les accroissements sont stationnaires, c'est à dire que, pour tout $t > 0$ et $h > 0$, la loi de la variable aléatoire $B(t + h) - B(t)$ ne dépend que de h.
4. Pour tout $t > 0$, $B(t) \sim \mathcal{N}(0, t)$

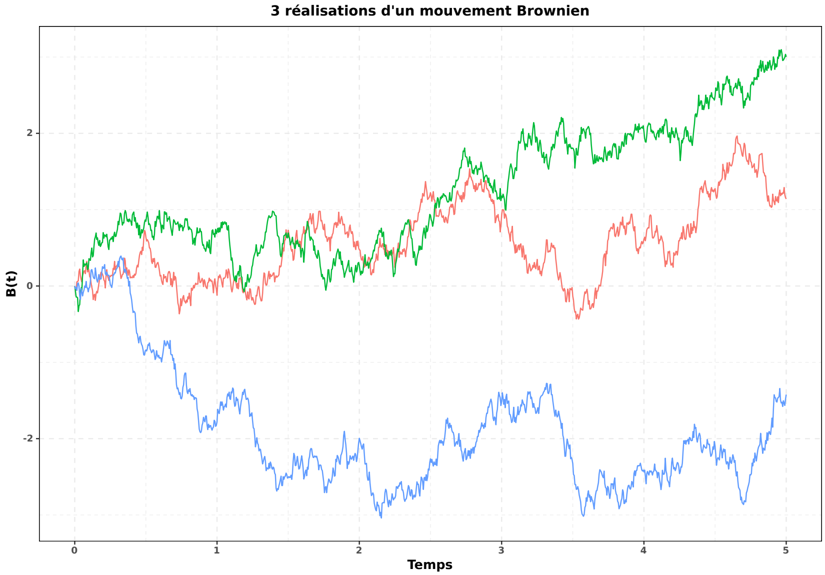
Remarque: Ne pas confondre $B(t)$: la variable aléatoire correspondant à la **valeur** du processus au temps t et $\{B(s)\}_{0\leq s\leq t}$: la collection (infinie et non dénombrable!) de V.A. correspondant à une **trajectoire** du processus jusqu'au temps t.

Propriétés

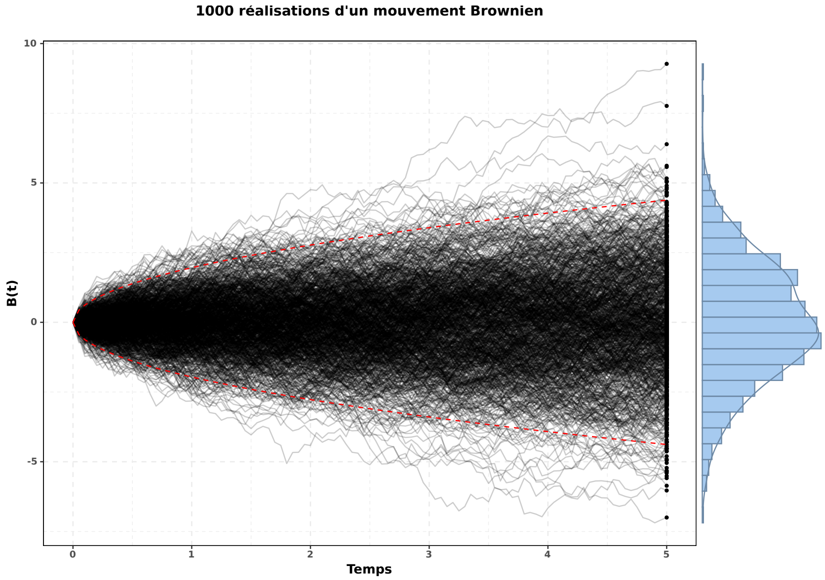
1. **Markov:** Le mouvement Brownien est un processus de Markov, c'est à dire que pour toute suite de temps $t_1 < \cdots < t_n$, la loi de $B(t_n)|B(t_1), \ldots, B(t_{n-1})$ est égale à la loi de $B(t_n)|B(t_{n-1})$.
2. **Loi de transition:**

$$B(t+\Delta)|B(t) \sim \mathcal{N}(B(t), \Delta).$$

Exemples de réalisations



Exemples de réalisations



Autres propriétés

On peut montrer que:

- **Continuité** Chaque trajectoire du mouvement Brownien est presque sûrement continue.
- **Non différentiabitlité** Le mouvement Brownien n'est dérivable en **aucun point** de sa trajectoire.

Autres propriétés

On peut montrer que:

- **Continuité** Chaque trajectoire du mouvement Brownien est presque sûrement continue.
- **Non différentiabitlité** Le mouvement Brownien n'est dérivable en **aucun point** de sa trajectoire.

Qualité comme modèle de bruit

1. **Accroissements indépendants** Pas d'effet d'accumulation des déviations autour de la vraie courbe

Autres propriétés

On peut montrer que:

- **Continuité** Chaque trajectoire du mouvement Brownien est presque sûrement continue.
- **Non différentiabitlité** Le mouvement Brownien n'est dérivable en **aucun point** de sa trajectoire.

Qualité comme modèle de bruit

1. **Accroissements indépendants** Pas d'effet d'accumulation des déviations autour de la vraie courbe
2. **Accroissements stationnaires** Pas de période spécifique où les déviations autour de la vraie courbe sont plus importantes.
3. **Accroissements de loi Normale** Déviations contrôlées.

Equations différentielles stochastiques

Modèle de croissance malthusienne

On considère une population de bactéries, partant d'une population x_0 et ayant un taux de reproduction $\alpha > 0$.

- L'accroissement infinitésimal de la population est donné par une proportion α de la population existante;

Modèle de croissance malthusienne

On considère une population de bactéries, partant d'une population x_0 et ayant un taux de reproduction $\alpha > 0$.

- L'accroissement infinitésimal de la population est donné par une proportion α de la population existante;
- Au temps t, la taille de la population satisfait l'équation différentielle:

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = \alpha x(t),\ x(0) = x_0$$

La solution de cette équation est bien connue:

$$x(t) = x_0 \exp(\alpha t)$$

Ajout de l'aléa

Dans l'environnement, il existe un aléa qui peut impacter l'accroissement infinitésimal, on a envie d'écrire:

$$\frac{\mathrm{d}X}{\mathrm{d}t}(t) = \alpha X(t) + \text{Aléa}, \; X(0) = x_0$$

La solution devient **aléatoire**, il s'agit d'un processus **stochastique**.

Ajout de l'aléa

Dans l'environnement, il existe un aléa qui peut impacter l'accroissement infinitésimal, on a envie d'écrire:

$$\frac{\mathrm{d}X}{\mathrm{d}t}(t) = \alpha X(t) + \text{Aléa}, \; X(0) = x_0$$

La solution devient **aléatoire**, il s'agit d'un processus **stochastique**.

Cela peut se réécrire

$$X(t) = X(0) + \int_0^t \alpha X(s)\mathrm{d}s + \text{Processus de perturbation.}$$

Ajout de l'aléa

Dans l'environnement, il existe un aléa qui peut impacter l'accroissement infinitésimal, on a envie d'écrire:

$$\frac{\mathrm{d}X}{\mathrm{d}t}(t) = \alpha X(t) + \text{Aléa}, \; X(0) = x_0$$

La solution devient **aléatoire**, il s'agit d'un processus **stochastique**.

Cela peut se réécrire

$$X(t) = X(0) + \int_0^t \alpha X(s)\mathrm{d}s + \text{Processus de perturbation.}$$

Ce processus de perturbation doit être modélisé. On choisira le **mouvement Brownien**.

1. **Accroissements indépendants** Pas d'effet d'accumulation des déviations autour de la vraie courbe.
2. **Accroissements stationnaires** Pas de période spécifique où les déviations autour de la vraie courbe sont plus importantes.
3. **Accroissements de loi Normale** Déviations contrôlées.

Equation différentielle stochastique

On pose donc:

$$X(t) = X(0) + \int_0^t \alpha X(s) \mathrm{d}s + B(t).$$

Equation différentielle stochastique

On pose donc:

$$X(t) = X(0) + \int_0^t \alpha X(s)\mathrm{d}s + B(t).$$

Pour ne pas oublier l'accumulation de la perturbation, on peut écrire:

$$X(t) = X(0) + \int_0^t \alpha X(s)\mathrm{d}s + \int_0^t \mathrm{d}B(s).$$

Equation différentielle stochastique

On pose donc:

$$X(t) = X(0) + \int_0^t \alpha X(s)\mathrm{d}s + B(t).$$

Pour ne pas oublier l'accumulation de la perturbation, on peut écrire:

$$X(t) = X(0) + \int_0^t \alpha X(s)\mathrm{d}s + \int_0^t \mathrm{d}B(s).$$

En pratique, on trouvera toujours l'écriture infinitésimale:

$$\mathrm{d}X(t) = \alpha X(t)\mathrm{d}t + \mathrm{d}B(t),\ X(0) = x_0$$

- Cette dernière écriture définit une **équation différentielle stochastique**.
- Sa solution est un **processus stochastique**.

Remarque fondamentale

La solution de

$$\mathrm{d}X(t) = \alpha X(t)\mathrm{d}t + \mathrm{d}B(t),\ X(0) = x_0$$

n'est pas

$$X(t) = x_0 \exp(\alpha t) + B(t).$$

et encore moins

$$X(t) = x_0 \exp(\alpha t) + \varepsilon(t),\ \varepsilon(t) \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$$

Remarque fondamentale

La solution de

$$\mathrm{d}X(t) = \alpha X(t)\mathrm{d}t + \mathrm{d}B(t),\ X(0) = x_0$$

n'est pas

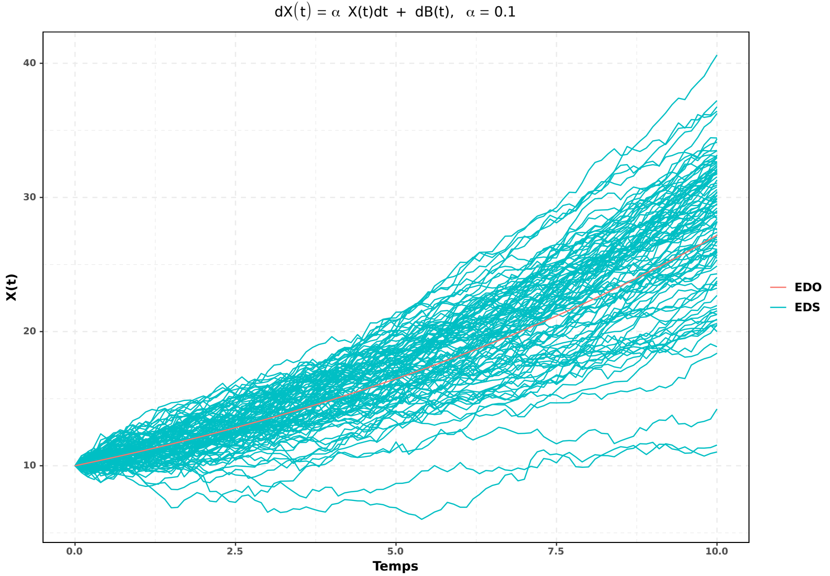
$$X(t) = x_0 \exp(\alpha t) + B(t).$$

et encore moins

$$X(t) = x_0 \exp(\alpha t) + \varepsilon(t),\ \varepsilon(t) \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$$

- Pour trouver une solution exacte, on doit passer par **le calcul stochastique.**
- En pratique, peu d'EDS ont une solution exacte (i.e. un processus dont la loi de transition est connue.)
- On passera par la simulation pour résoudre les EDS.

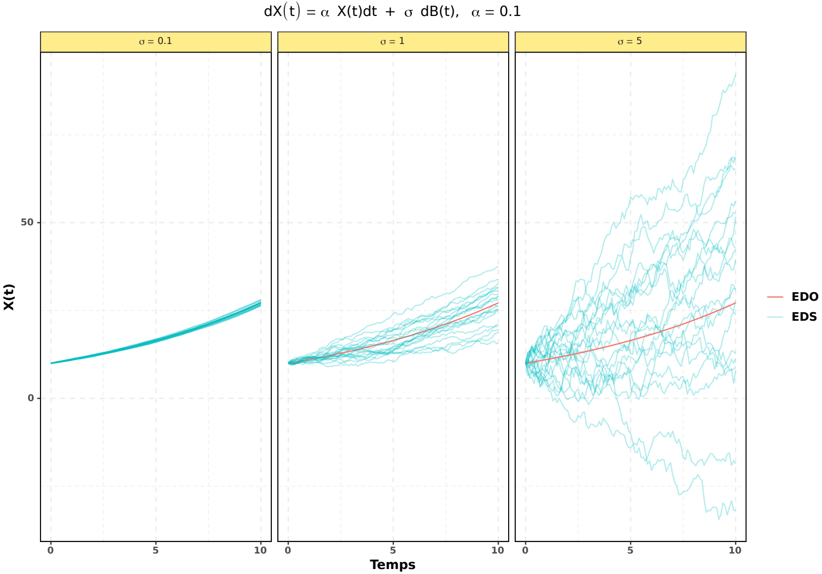
Exemple (1)



$dX(t) = \alpha\ X(t)dt\ +\ dB(t),\ \ \alpha = 0.1$
40
30
20
10
X(t)
0.0
2.5
5.0
7.5
10.0
Temps
EDO
EDS

Exemple (2), amplification de la perturbation

On peut ajouter un paramètre d'importance de la perturbation σ.



Complexification du modèle

Pour éviter des possibles valeurs négatives, on veut que les perturbations soient faibles quand la population est petite, on pose:

$$\mathrm{d}X(t) = \alpha X(t)\mathrm{d}t + \sigma X(t)\mathrm{d}B(t),\ X(0) = x_0$$

Complexification du modèle

Pour éviter des possibles valeurs négatives, on veut que les perturbations soient faibles quand la population est petite, on pose:

$$\mathrm{d}X(t) = \alpha X(t)\mathrm{d}t + \sigma X(t)\mathrm{d}B(t),\ X(0) = x_0$$

La solution est alors donnée par:

$$X(t) = X(0) + \int_0^t \alpha X(s)\mathrm{d}s + \int_0^t \sigma X(s)\mathrm{d}B(s).$$

Complexification du modèle

Pour éviter des possibles valeurs négatives, on veut que les perturbations soient faibles quand la population est petite, on pose:

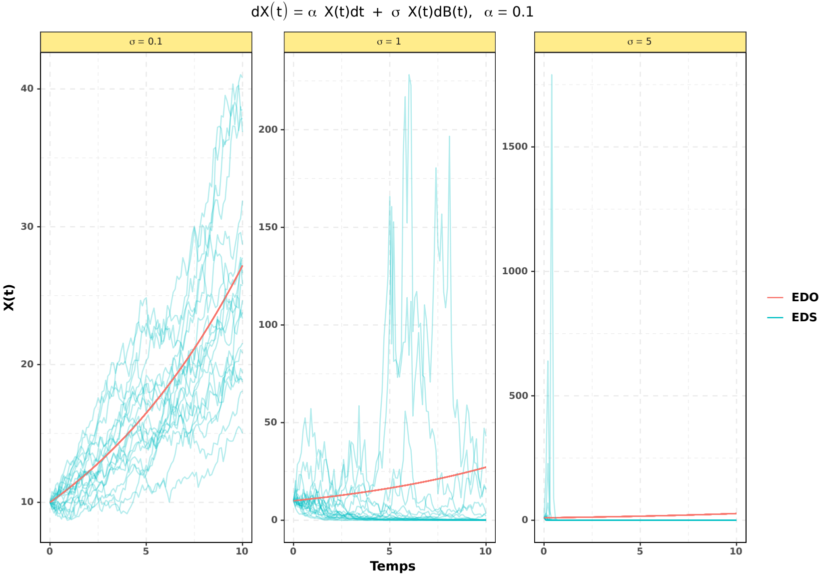
$$\mathrm{d}X(t) = \alpha X(t)\mathrm{d}t + \sigma X(t)\mathrm{d}B(t),\ X(0) = x_0$$

La solution est alors donnée par:

$$X(t) = X(0) + \int_0^t \alpha X(s)\mathrm{d}s + \int_0^t \sigma X(s)\mathrm{d}B(s).$$

Remarque: La dernière intégrale n'a pas un sens évident. Elle est proprement définie par le **calcul stochastique** (ou calcul d'Ito).

Exemple (3)



Modèle générique d'équation différentielle stochastique

De manière générique, une EDS s'écrit:

$$\mathrm{d}X(t) = f(X(t), t)\mathrm{d}t + g(X(t), t)\mathrm{d}B(t),\ X(0) \sim \chi_0$$

- χ_0 est la **distribution initiale**;
- $f()$ est la **fonction de dérive** (*drift*);
- $g()$ est la **fonction de diffusion**.

Modèle générique d'équation différentielle stochastique

De manière générique, une EDS s'écrit:

$$\mathrm{d}X(t) = f(X(t), t)\mathrm{d}t + g(X(t), t)\mathrm{d}B(t),\ X(0) \sim \chi_0$$

- χ_0 est la **distribution initiale**;
- $f()$ est la **fonction de dérive** (*drift*);
- $g()$ est la **fonction de diffusion**.

La solution est alors donnée par:

$$X(t) = X(0) + \int_0^t f(X(s), s)\mathrm{d}s + \int_0^t g(X(s), s)\mathrm{d}B(s).$$

Loi de transtion et simulation

Pour simuler, il faut connaître la loi de $X(t)|X(0)$;

L'algorithme est alors:

1. Tirer $X(0)$ selon la loi initiale χ_0;
2. Tirer $X(t)$ selon la loi de $X(t)|X(0)$.

Loi de transtion et simulation

Pour simuler, il faut connaître la loi de $X(t)|X(0)$;

L'algorithme est alors:

1. Tirer $X(0)$ selon la loi initiale χ_0;
2. Tirer $X(t)$ selon la loi de $X(t)|X(0)$.

Problème, pour la plupart des EDS, cette loi est **inconnue**!

Loi de transtion et simulation

Pour simuler, il faut connaître la loi de $X(t)|X(0)$;

L'algorithme est alors:

1. Tirer $X(0)$ selon la loi initiale χ_0;
2. Tirer $X(t)$ selon la loi de $X(t)|X(0)$.

Problème, pour la plupart des EDS, cette loi est **inconnue**!On peut recourir à une **approximation**.

Approximation d'Euler

La solution est donnée par:

$$X(t) = X(0) + \int_0^t f(X(s), s)\mathrm{d}s + \int_0^t g(X(s), s)\mathrm{d}B(s).$$

Approximation d'Euler

La solution est donnée par:

$$X(t) = X(0) + \int_0^t f(X(s), s)\mathrm{d}s + \int_0^t g(X(s), s)\mathrm{d}B(s).$$

Pour h petit, on a:

$$X(h) \approx X(0) + f(X(0), 0) \times h + g(X(0), 0) \times \overbrace{(B(h) - B(0))}^{\sim \mathcal{N}(0,h)}$$

Approximation d'Euler

La solution est donnée par:

$$X(t) = X(0) + \int_0^t f(X(s), s)\mathrm{d}s + \int_0^t g(X(s), s)\mathrm{d}B(s).$$

Pour h petit, on a:

$$X(h) \approx X(0) + f(X(0), 0) \times h + g(X(0), 0) \times \overbrace{(B(h) - B(0))}^{\sim \mathcal{N}(0,h)}$$

Donc, pour h petit, on a une **approximation de la loi de transition**:

$$X(h)|X(0) \overset{\approx}{\sim} \mathcal{N}\left(X(0) + f(X(0), 0) \times h, g(X(0), 0)^2 \times h\right)$$

Cettte approximation est appelée **schéma d'Euler**.

Pseudo algorithme

$$\mathrm{d}X(t) = f(X(t), t)\mathrm{d}t + g(X(t), t)\mathrm{d}B(t),\ X(0) \sim \chi_0$$

On peut donc simuler à un vecteur de temps $(0, t_1 = h, t_2 = 2h, \ldots, t_n = nh)$.

1. **Initialisation:** On obtient x_0 en simulant selon χ_0 (ou on le fixe);

Pseudo algorithme

$$\mathrm{d}X(t) = f(X(t), t)\mathrm{d}t + g(X(t), t)\mathrm{d}B(t),\ X(0) \sim \chi_0$$

On peut donc simuler à un vecteur de temps $(0, t_1 = h, t_2 = 2h, \ldots, t_n = nh)$.

1. **Initialisation:** On obtient x_0 en simulant selon χ_0 (ou on le fixe);
2. **Itération:** Pour k allant de 1 à n:
   - On obtient x_k en tirant dans une loi normale

$$\mathcal{N}\left(x_{k-1} + f(x_{k-1}, t_{k-1}) \times h, g(x_{k-1}, t_{k-1})^2 \times h\right)$$

Squelette de code de simulation en R

$$\mathrm{d}X(t) = f(X(t), t)\mathrm{d}t + g(X(t), t)\mathrm{d}B(t),\ \ X(0) \sim \chi_0$$

```
# On suppose qu'on a déjà codé les fonctions f(x, t) et g(x, t)
# x0 a été obtenu (ou choisi) préalablement
# times est le vecteur de temps auxquels on simule
simulate_sde <- function(x0, times, ...){
  n_points <- length(times) # Nombre de points de simulations
  output <- rep(NA, n_points) # Initialisation du vecteur final
  output[1] <- x0 # Initialisation
  for(k in 2:n_points){ # Itération
    h <- times[k] - times[k - 1] # Pas de temps (doit être petit!)
    moyenne_euler <- output[k - 1] + # x
      f(output[k - 1], times[k - 1]) * h # f(x, t) * h
    variance_euler <- g(output[k - 1], times[k - 1])^2 * h
    output[k] <- rnorm(n = 1, # 1 simulation de loi normale
                       mean = moyenne_euler, # Moyenne
                       sd = sqrt(variance_euler) # Ecart-type
                       )
  }
  return(output)
}
```