

Processus stochastiques

Un processus stochastique est une famille $Y = \{Y_t, t \in I\}$ de variables aléatoires définies sur un même espace de probabilité.

Processus stochastiques

Un **processus stochastique** est une famille $Y = \{Y_t, t \in I\}$ de variables aléatoires définies sur un même espace de probabilité.

L'indice t est souvent interprété comme le **temps**.

Le processus est en **temps continu** si I est continu (e.g., $I = [0, \infty)$),
et en **temps discret** si I est discret (e.g., $I = \{0, 1, 2, \dots\}$).

Processus stochastiques

Un **processus stochastique** est une famille $Y = \{Y_t, t \in I\}$ de variables aléatoires définies sur un même espace de probabilité.

L'indice t est souvent interprété comme le **temps**.

Le processus est en **temps continu** si I est continu (e.g., $I = [0, \infty)$), et en **temps discret** si I est discret (e.g., $I = \{0, 1, 2, \dots\}$).

Lorsque t est continu, on note souvent Y_t par $Y(t)$.

Processus stochastiques

Un **processus stochastique** est une famille $Y = \{Y_t, t \in I\}$ de variables aléatoires définies sur un même espace de probabilité.

L'indice t est souvent interprété comme le **temps**.

Le processus est en **temps continu** si I est continu (e.g., $I = [0, \infty)$), et en **temps discret** si I est discret (e.g., $I = \{0, 1, 2, \dots\}$).

Lorsque t est continu, on note souvent Y_t par $Y(t)$.

On supposera ici que Y_t prend ses valeurs dans \mathbb{R}^d .

Processus markoviens

Un processus stochastique est **Markovien** si, conditionnellement à sa valeur présente au temps t , son évolution future est indépendante de son passé.

Processus markoviens

Un processus stochastique est **Markovien** si, conditionnellement à sa valeur présente au temps t , son évolution future est indépendante de son passé.

C'est-à-dire: pour toute variable aléatoire X fonction de $\{Y_s, s > t\}$, la loi de X conditionnelle à $\{Y_s, s \leq t\}$, est la même que celle conditionnelle à Y_t .

Processus markoviens

Un processus stochastique est **Markovien** si, conditionnellement à sa valeur présente au temps t , son évolution future est indépendante de son passé.

C'est-à-dire: pour toute variable aléatoire X fonction de $\{Y_s, s > t\}$, la loi de X conditionnelle à $\{Y_s, s \leq t\}$, est la même que celle conditionnelle à Y_t .

Veut dire que Y_t contient toujours assez d'information pour "générer" le futur.
(Suffit de mettre assez d'information dans Y_t .)

Processus markoviens

Un processus stochastique est **Markovien** si, conditionnellement à sa valeur présente au temps t , son évolution future est indépendante de son passé.

C'est-à-dire: pour toute variable aléatoire X fonction de $\{Y_s, s > t\}$, la loi de X conditionnelle à $\{Y_s, s \leq t\}$, est la même que celle conditionnelle à Y_t .

Veut dire que Y_t contient toujours assez d'information pour "générer" le futur.
(Suffit de mettre assez d'information dans Y_t .)

Chaîne de Markov temps discret: lorsque $I = \{0, 1, \dots\}$.

Exemple: File $GI/GI/1$.

$\{W_i, i \geq 0\}$ est markovien, car (Eq. de Lindley) $W_{i+1} = \max(0, W_i + S_i - A_i)$.

Exemple: File $GI/GI/1$.

$\{W_i, i \geq 0\}$ est markovien, car (Eq. de Lindley) $W_{i+1} = \max(0, W_i + S_i - A_i)$.

$\{Q(t), t \geq 0\}$ n'est pas markovien, sauf pour une file $M/M/1$.

Exemple: File $GI/GI/1$.

$\{W_i, i \geq 0\}$ est markovien, car (Eq. de Lindley) $W_{i+1} = \max(0, W_i + S_i - A_i)$.

$\{Q(t), t \geq 0\}$ n'est pas markovien, sauf pour une file $M/M/1$.

Posons: T = horizon de la simulation;

t_i = instant du i -ième événement e_i ;

Q_i = nombre de clients dans la file juste après e_i ;

ζ_i = temps jusqu'à la prochaine arrivée;

ξ_i = durée résiduelle de service (-1 si personne);

$\mathcal{S}_i = (t_i, Q_i, \zeta_i, \xi_i)$ = état de la simulation.

Exemple: File $GI/GI/1$.

$\{W_i, i \geq 0\}$ est markovien, car (Eq. de Lindley) $W_{i+1} = \max(0, W_i + S_i - A_i)$.

$\{Q(t), t \geq 0\}$ n'est pas markovien, sauf pour une file $M/M/1$.

Posons: T = horizon de la simulation;

t_i = instant du i -ième événement e_i ;

Q_i = nombre de clients dans la file juste après e_i ;

ζ_i = temps jusqu'à la prochaine arrivée;

ξ_i = durée résiduelle de service (-1 si personne);

$\mathcal{S}_i = (t_i, Q_i, \zeta_i, \xi_i)$ = état de la simulation.

Le processus $\mathcal{S} = \{\mathcal{S}_i, i \geq 0\}$ est une chaîne de Markov.

En général, $\{\mathcal{S}_i, i = 0, 1, 2, \dots\}$, où \mathcal{S}_i est l'état du modèle de simulation au temps t_i , juste après l'événement e_i , est une chaîne de Markov.

En général, $\{\mathcal{S}_i, i = 0, 1, 2, \dots\}$, où \mathcal{S}_i est l'état du modèle de simulation au temps t_i , juste après l'événement e_i , est une chaîne de Markov.

Par contre, le processus $\{\mathcal{S}(t), t \geq 0\}$ défini par $\mathcal{S}(t) = \mathcal{S}_{N(t)}$, où $N(t) = \sup\{i \mid t_i \leq t\}$ est le nombre d'événements durant $(0, t]$, ne l'est pas.

Le processus $\{\mathcal{S}(t), t \geq 0\}$ est uniquement déterminé par $\{\mathcal{S}_i, i = 0, 1, \dots\}$, mais pas l'inverse. Il peut y avoir des événements simultanés.

En général, $\{\mathcal{S}_i, i = 0, 1, 2, \dots\}$, où \mathcal{S}_i est l'état du modèle de simulation au temps t_i , juste après l'événement e_i , est une chaîne de Markov.

Par contre, le processus $\{\mathcal{S}(t), t \geq 0\}$ défini par $\mathcal{S}(t) = \mathcal{S}_{N(t)}$, où $N(t) = \sup\{i \mid t_i \leq t\}$ est le nombre d'événements durant $(0, t]$, ne l'est pas.

Le processus $\{\mathcal{S}(t), t \geq 0\}$ est uniquement déterminé par $\{\mathcal{S}_i, i = 0, 1, \dots\}$, mais pas l'inverse. Il peut y avoir des événements simultanés.

Pour les chaînes de Markov, on peut calculer les mesures de performance **numériquement** si le nombre d'états n'est pas trop grand.

Filtrages et temps d'arrêt

Pour un processus $Y = \{Y_i, i = 0, 1, \dots\}$, $\mathcal{F}_t = \sigma(Y_0, \dots, Y_t)$ représente toute l'information que l'on peut déduire en observant le processus jusqu'à l'étape t .

Filtrages et temps d'arrêt

Pour un processus $Y = \{Y_i, i = 0, 1, \dots\}$, $\mathcal{F}_t = \sigma(Y_0, \dots, Y_t)$ représente toute l'information que l'on peut déduire en observant le processus jusqu'à l'étape t .

De même, pour $Y = \{Y_t, t \geq 0\}$, \mathcal{F}_t représente l'information que l'on peut déduire de $\{Y_s, 0 \leq s \leq t\}$.

La famille $\{\mathcal{F}_t, t \geq 0\}$ s'appelle un **filtrage** (généré par Y).

Filtrages et temps d'arrêt

Pour un processus $Y = \{Y_i, i = 0, 1, \dots\}$, $\mathcal{F}_t = \sigma(Y_0, \dots, Y_t)$ représente toute l'information que l'on peut déduire en observant le processus jusqu'à l'étape t .

De même, pour $Y = \{Y_t, t \geq 0\}$, \mathcal{F}_t représente l'information que l'on peut déduire de $\{Y_s, 0 \leq s \leq t\}$.

La famille $\{\mathcal{F}_t, t \geq 0\}$ s'appelle un **filtrage** (généré par Y).

Une variable aléatoire est **\mathcal{F}_t -mesurable** si on peut toujours déduire sa valeur à partir de \mathcal{F}_t .

Un **temps d'arrêt** par rapport à $\{\mathcal{F}_t, t \geq 0\}$ est une variable aléatoire T , à valeurs dans I , telle que T est \mathcal{F}_T -mesurable.

Exemple. Centre d'appels.

Soit $\{\mathcal{F}_i, i \geq 0\}$ le filtrage générée par $\{\mathcal{S}_i, i \geq 0\}$.

N = nombre d'événements durant la journée.

N' = numéro d'événement de la dernière arrivée.

N est un temps d'arrêt par rapport à $\{\mathcal{F}_i, i \geq 0\}$, mais pas N' , sauf si $\mathcal{S}_{N'}$ contient l'information requise pour que l'on sache que c'est la dernière arrivée.

Marche aléatoire

Processus $\{S_j, j \geq 0\}$ défini sur \mathbb{R} par

$$S_j = S_{j-1} + X_j,$$

où $S_0 = s_0$ (une constante) et les X_j sont des v.a. i.i.d..

Généralisation: Marche aléatoire en d -dimensions, dans \mathbb{R}^d .

On approxime souvent les processus en temps continu par des marches aléatoires, pour pouvoir simuler les trajectoires.

Processus de Poisson

Un **processus de comptage** est un processus en temps continu $\{N(t), t \geq 0\}$, à valeurs dans $\{0, 1, 2, \dots\}$, et dont les trajectoires sont non décroissantes et continues à droite.

Habituellement, $N(0) = 0$, et les instants de saut $0 < T_1 \leq T_2 \leq \dots \leq T_j \leq \dots$ s'appellent les instants d'**arrivées**.

Ainsi, $N(t)$ représente le nombre d'arrivées durant $[0, t]$.

Processus de Poisson

Un **processus de comptage** est un processus en temps continu $\{N(t), t \geq 0\}$, à valeurs dans $\{0, 1, 2, \dots\}$, et dont les trajectoires sont non décroissantes et continues à droite.

Habituellement, $N(0) = 0$, et les instants de saut $0 < T_1 \leq T_2 \leq \dots \leq T_j \leq \dots$ s'appellent les instants d'**arrivées**.

Ainsi, $N(t)$ représente le nombre d'arrivées durant $[0, t]$.

On note $A_j = T_j - T_{j-1}$.

Si les A_j sont des v.a. i.i.d., on a un **processus de renouvellement**.

Processus de Poisson

Un **processus de comptage** est un processus en temps continu $\{N(t), t \geq 0\}$, à valeurs dans $\{0, 1, 2, \dots\}$, et dont les trajectoires sont non décroissantes et continues à droite.

Habituellement, $N(0) = 0$, et les instants de saut $0 < T_1 \leq T_2 \leq \dots \leq T_j \leq \dots$ s'appellent les instants d'**arrivées**.

Ainsi, $N(t)$ représente le nombre d'arrivées durant $[0, t]$.

On note $A_j = T_j - T_{j-1}$.

Si les A_j sont des v.a. i.i.d., on a un **processus de renouvellement**.

On a un **processus de Poisson** si $N(0) = 0$ et

- (a) les arrivées se font une à une (la prob. de 2 arrivées simultanées est 0);
- (b) pour $s, t \geq 0$, la v.a. $N(t + s) - N(t)$ est indépendante de $\{N(u), u \leq t\}$ (i.e., ne dépend pas de l'histoire passée).

Processus de Poisson

Un **processus de comptage** est un processus en temps continu $\{N(t), t \geq 0\}$, à valeurs dans $\{0, 1, 2, \dots\}$, et dont les trajectoires sont non décroissantes et continues à droite.

Habituellement, $N(0) = 0$, et les instants de saut $0 < T_1 \leq T_2 \leq \dots \leq T_j \leq \dots$ s'appellent les instants d'**arrivées**.

Ainsi, $N(t)$ représente le nombre d'arrivées durant $[0, t]$.

On note $A_j = T_j - T_{j-1}$.

Si les A_j sont des v.a. i.i.d., on a un **processus de renouvellement**.

On a un **processus de Poisson** si $N(0) = 0$ et

- (a) les arrivées se font une à une (la prob. de 2 arrivées simultanées est 0);
- (b) pour $s, t \geq 0$, la v.a. $N(t + s) - N(t)$ est indépendante de $\{N(u), u \leq t\}$ (i.e., ne dépend pas de l'histoire passée).

Intuition: un processus sera approx. Poisson si les événements arrivent au hasard, indépendamment les uns des autres.

Posons $a(t) \stackrel{\text{def}}{=} \mathbb{E}[N(t)]$.

Supposons que $a(t)$ est continue partout, et dérivable sauf possiblement en un nombre fini de points sur tout intervalle fini.

Posons $a(t) \stackrel{\text{def}}{=} \mathbb{E}[N(t)]$.

Supposons que $a(t)$ est continue partout, et dérivable sauf possiblement en un nombre fini de points sur tout intervalle fini.

$\lambda(t) = a'(t)$ est la **fonction de taux** du processus;

$a(t) = \int_0^t \lambda(s) ds$ est la **fonction de taux cumulé**.

Posons $a(t) \stackrel{\text{def}}{=} \mathbb{E}[N(t)]$.

Supposons que $a(t)$ est continue partout, et dérivable sauf possiblement en un nombre fini de points sur tout intervalle fini.

$\lambda(t) = a'(t)$ est la **fonction de taux** du processus;

$a(t) = \int_0^t \lambda(s)ds$ est la **fonction de taux cumulé**.

Interprétation: pour un petit $\epsilon > 0$, la probabilité d'un saut du processus dans l'intervalle de temps $(t, t + \epsilon]$ est

$$\begin{aligned} P[N(t + \epsilon) - N(t) = 1] &\approx 1 - P[N(t + \epsilon) - N(t) = 0] \\ &\approx \mathbb{E}[N(t + \epsilon)] - \mathbb{E}[N(t)] \approx \lambda(t)\epsilon \end{aligned}$$

et la probabilité de plus d'un saut est $o(\epsilon)$.

Posons $a(t) \stackrel{\text{def}}{=} \mathbb{E}[N(t)]$.

Supposons que $a(t)$ est continue partout, et dérivable sauf possiblement en un nombre fini de points sur tout intervalle fini.

$\lambda(t) = a'(t)$ est la **fonction de taux** du processus;

$a(t) = \int_0^t \lambda(s) ds$ est la **fonction de taux cumulé**.

Interprétation: pour un petit $\epsilon > 0$, la probabilité d'un saut du processus dans l'intervalle de temps $(t, t + \epsilon]$ est

$$\begin{aligned} P[N(t + \epsilon) - N(t) = 1] &\approx 1 - P[N(t + \epsilon) - N(t) = 0] \\ &\approx \mathbb{E}[N(t + \epsilon)] - \mathbb{E}[N(t)] \approx \lambda(t)\epsilon \end{aligned}$$

et la probabilité de plus d'un saut est $o(\epsilon)$.

Processus de Poisson **stationnaire**: $\lambda(t) = \lambda > 0$ pour tout $t \geq 0$.

Si $\lambda = 1$, on a un processus de Poisson **standard**.

Proposition. Si $\{N(t), t \geq 0\}$ est un processus de Poisson et $t_2 > t_1 \geq 0$, alors $N(t_2) - N(t_1)$ suit la loi de Poisson de moyenne

$$a(t_2) - a(t_1) = \int_{t_1}^{t_2} \lambda(t) dt.$$

Proposition. Si $\{N(t), t \geq 0\}$ est un processus de Poisson et $t_2 > t_1 \geq 0$, alors $N(t_2) - N(t_1)$ suit la loi de Poisson de moyenne

$$a(t_2) - a(t_1) = \int_{t_1}^{t_2} \lambda(t) dt.$$

Dans le cas stationnaire, la moyenne est $(t_2 - t_1)\lambda$.

Proposition. Si $\{N(t), t \geq 0\}$ est un processus de Poisson et $t_2 > t_1 \geq 0$, alors $N(t_2) - N(t_1)$ suit la loi de Poisson de moyenne

$$a(t_2) - a(t_1) = \int_{t_1}^{t_2} \lambda(t) dt.$$

Dans le cas stationnaire, la moyenne est $(t_2 - t_1)\lambda$.

Proposition. Un processus de comptage $\{N(t), t \geq 0\}$, avec $N(0) = 0$, est un processus de Poisson stationnaire de taux λ ssi les v.a. A_1, A_2, \dots sont i.i.d. exponentielles de paramètre λ .

Proposition. Si $\{N(t), t \geq 0\}$ est un processus de Poisson et $t_2 > t_1 \geq 0$, alors $N(t_2) - N(t_1)$ suit la loi de Poisson de moyenne

$$a(t_2) - a(t_1) = \int_{t_1}^{t_2} \lambda(t) dt.$$

Dans le cas stationnaire, la moyenne est $(t_2 - t_1)\lambda$.

Proposition. Un processus de comptage $\{N(t), t \geq 0\}$, avec $N(0) = 0$, est un processus de Poisson stationnaire de taux λ ssi les v.a. A_1, A_2, \dots sont i.i.d. exponentielles de paramètre λ .

Dans le cas stationnaire, on peut donc générer les sauts en générant des exponentielles i.i.d..

On sait que pour un processus de Poisson, le **nombre d'arrivées** durant un intervalle $(t_1, t_2]$ est $\text{Poisson}(a(t_2) - a(t_1))$. On peut donc générer ce nombre directement. ¹¹

On sait que pour un processus de Poisson, le **nombre d'arrivées** durant un intervalle $(t_1, t_2]$ est $\text{Poisson}(a(t_2) - a(t_1))$. On peut donc générer ce nombre directement. Mais comment générer ensuite les **instants** de ces arrivées?

On sait que pour un processus de Poisson, le **nombre d'arrivées** durant un intervalle $(t_1, t_2]$ est $\text{Poisson}(a(t_2) - a(t_1))$. On peut donc générer ce nombre directement.

Mais comment générer ensuite les **instants** de ces arrivées?

Dans le cas **stationnaire**, c'est facile:

Proposition. Pour un processus de Poisson **stationnaire**, si n arrivées ont eu lieu durant un intervalle $(t_1, t_2]$, alors la loi conditionnelle des temps de ces n arrivées est la même que celle des statistiques d'ordre de n v.a. i.i.d. $U(t_1, t_2)$.

On sait que pour un processus de Poisson, le **nombre d'arrivées** durant un intervalle $(t_1, t_2]$ est $\text{Poisson}(a(t_2) - a(t_1))$. On peut donc générer ce nombre directement.

Mais comment générer ensuite les **instants** de ces arrivées?

Dans le cas **stationnaire**, c'est facile:

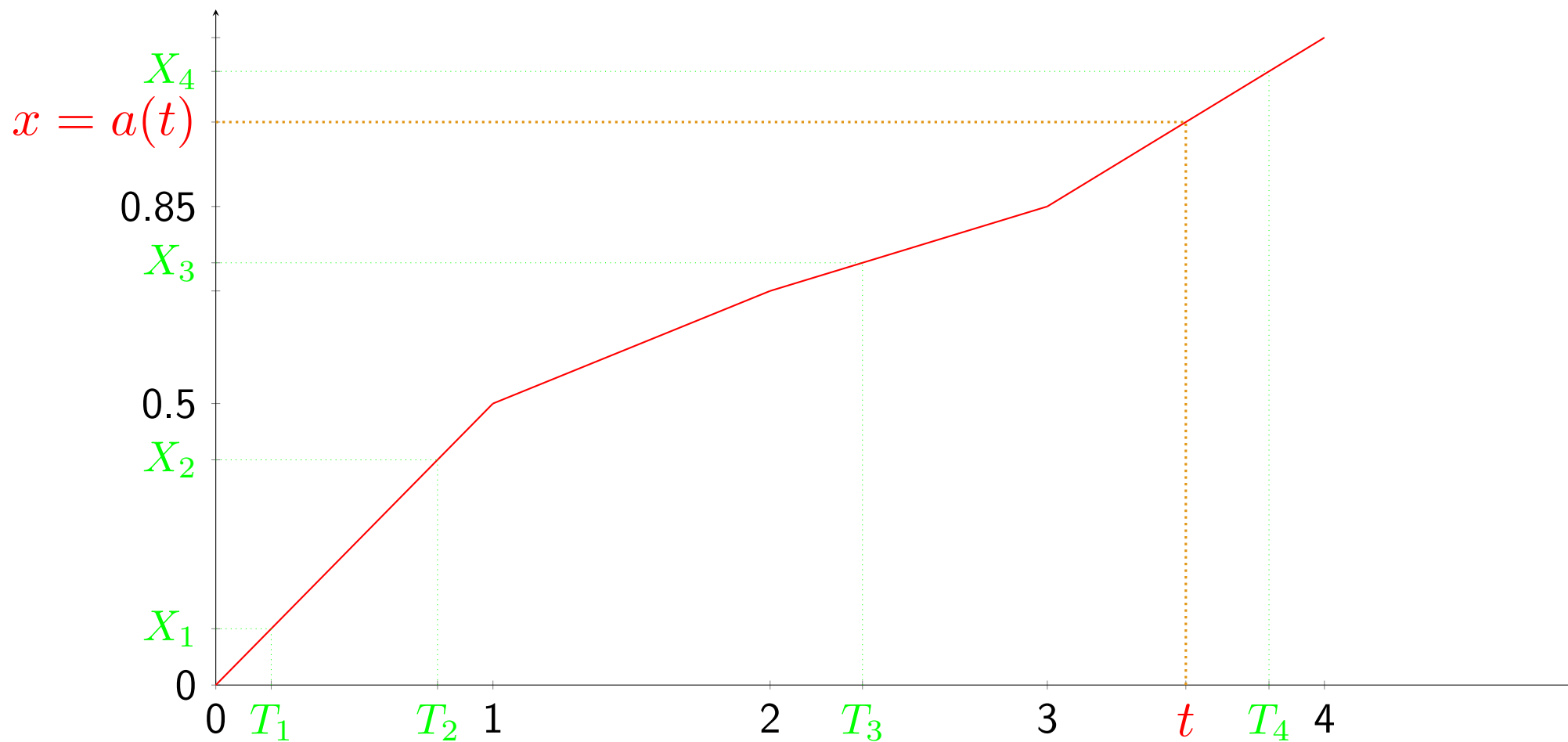
Proposition. Pour un processus de Poisson **stationnaire**, si n arrivées ont eu lieu durant un intervalle $(t_1, t_2]$, alors la loi conditionnelle des temps de ces n arrivées est la même que celle des statistiques d'ordre de n v.a. i.i.d. $U(t_1, t_2)$.

Encore une autre façon: “**Poisson bridge sampling**”, en exploitant le fait que le processus de Poisson stationnaire est un processus de Lévy.

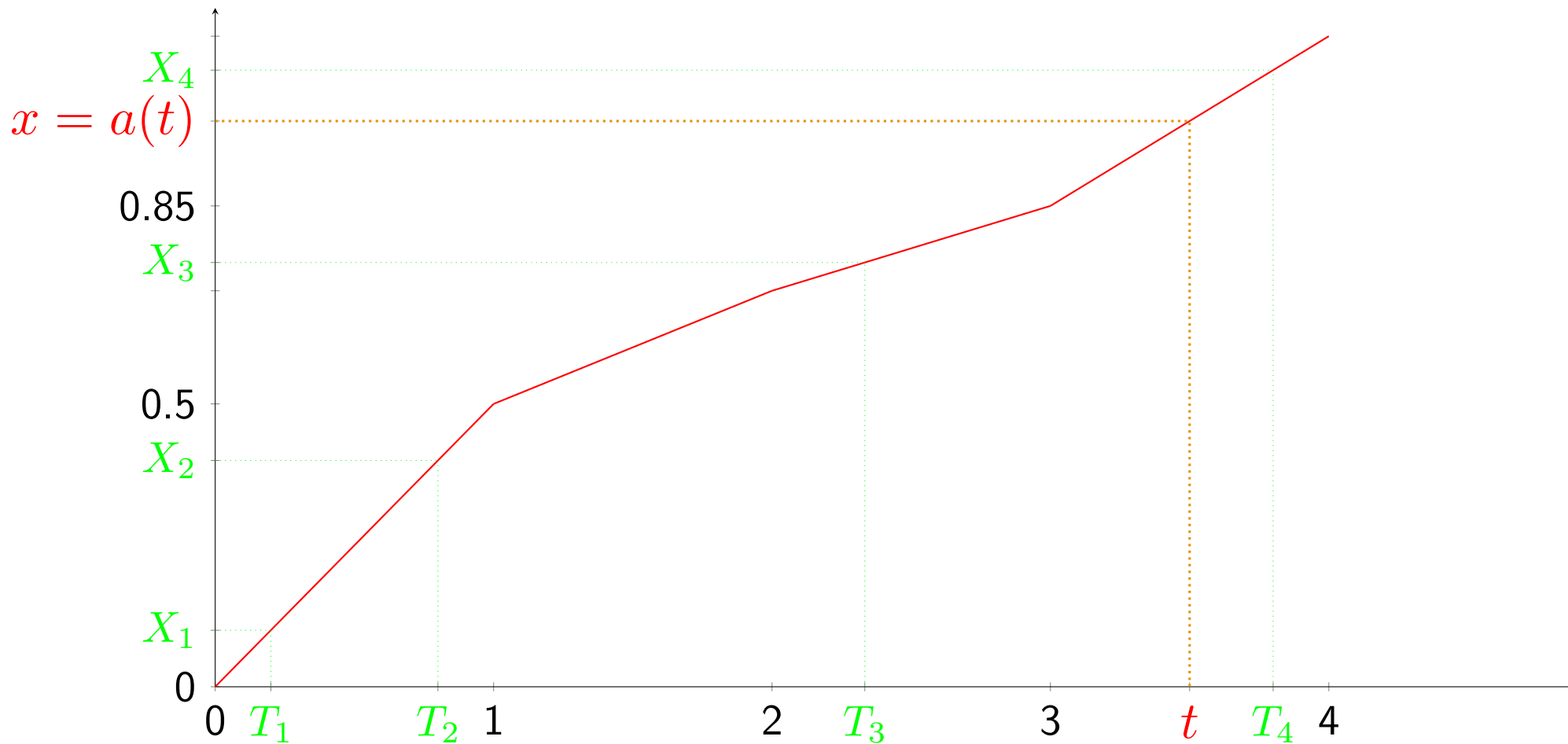
Pour l'intervalle $(0, t]$: générer d'abord $N(t)$ (Poisson);

puis $N(t/2)$ conditionnement à $N(t)$ (**binomiale**);

puis $N(t/4)$ conditionnement à $N(t/2)$ (binomiale); etc.



Ici, pour $t = 3.5$, on a $x = a(t) = 1.0$ et $N(3.5) = N_0(a(3.5)) = N_0(1) = 3$.



Ici, pour $t = 3.5$, on a $x = a(t) = 1.0$ et $N(3.5) = N_0(a(3.5)) = N_0(1) = 3$.

Note: Si $\lambda(\cdot)$ est constant [par morceaux], alors $a(\cdot)$ est linéaire [par morceaux].

Si $\lambda(\cdot)$ est linéaire, alors $a(\cdot)$ est quadratique (plus coûteux à inverser).

Standardisation

Un processus non-stationnaire peut être transformé en processus stationnaire simplement en étirant et comprimant l'échelle du temps.

Standardisation

Un processus non-stationnaire peut être transformé en processus stationnaire simplement en étirant et comprimant l'échelle du temps.

Supposons que l'on veut un processus de Poisson de taux $\lambda(t)$ et de taux cumulé $a(t)$ au temps t , pour tout $t \geq 0$.

Soit $N_0 = \{N_0(x), x \geq 0\}$ un processus de Poisson standard (de taux 1) et définissons le processus $N = \{N(t), t \geq 0\}$ via $N(t) = N_0(a(t))$ pour $t \geq 0$. Si les sauts de N sont aux temps T_1, T_2, \dots et ceux de N_0 sont à X_1, X_2, \dots , alors on a $X_j = a(T_j)$, ou encore $T_j = a^{-1}(X_j)$.

Standardisation

Un processus non-stationnaire peut être transformé en processus stationnaire simplement en étirant et comprimant l'échelle du temps.

Supposons que l'on veut un processus de Poisson de taux $\lambda(t)$ et de taux cumulé $a(t)$ au temps t , pour tout $t \geq 0$.

Soit $N_0 = \{N_0(x), x \geq 0\}$ un processus de Poisson standard (de taux 1) et définissons le processus $N = \{N(t), t \geq 0\}$ via $N(t) = N_0(a(t))$ pour $t \geq 0$. Si les sauts de N sont aux temps T_1, T_2, \dots et ceux de N_0 sont à X_1, X_2, \dots , alors on a $X_j = a(T_j)$, ou encore $T_j = a^{-1}(X_j)$.

Proposition $\{N(t), t \geq 0\}$ est un processus de Poisson de taux cumulé $a(\cdot)$ ssi $\{N_0(x), x \geq 0\}$ est un processus de Poisson standard.

Standardisation

Un processus non-stationnaire peut être transformé en processus stationnaire simplement en étirant et comprimant l'échelle du temps.

Supposons que l'on veut un processus de Poisson de taux $\lambda(t)$ et de taux cumulé $a(t)$ au temps t , pour tout $t \geq 0$.

Soit $N_0 = \{N_0(x), x \geq 0\}$ un processus de Poisson standard (de taux 1) et définissons le processus $N = \{N(t), t \geq 0\}$ via $N(t) = N_0(a(t))$ pour $t \geq 0$. Si les sauts de N sont aux temps T_1, T_2, \dots et ceux de N_0 sont à X_1, X_2, \dots , alors on a $X_j = a(T_j)$, ou encore $T_j = a^{-1}(X_j)$.

Proposition $\{N(t), t \geq 0\}$ est un processus de Poisson de taux cumulé $a(\cdot)$ ssi $\{N_0(x), x \geq 0\}$ est un processus de Poisson standard.

Preuve. Les axiomes (a) et (b) pour un processus de Poisson sont satisfaites pour N ssi elles le sont pour N_0 , car $a(\cdot)$ est continue et non-décroissante.

Standardisation

Un processus non-stationnaire peut être transformé en processus stationnaire simplement en étirant et comprimant l'échelle du temps.

Supposons que l'on veut un processus de Poisson de taux $\lambda(t)$ et de taux cumulé $a(t)$ au temps t , pour tout $t \geq 0$.

Soit $N_0 = \{N_0(x), x \geq 0\}$ un processus de Poisson standard (de taux 1) et définissons le processus $N = \{N(t), t \geq 0\}$ via $N(t) = N_0(a(t))$ pour $t \geq 0$. Si les sauts de N sont aux temps T_1, T_2, \dots et ceux de N_0 sont à X_1, X_2, \dots , alors on a $X_j = a(T_j)$, ou encore $T_j = a^{-1}(X_j)$.

Proposition $\{N(t), t \geq 0\}$ est un processus de Poisson de taux cumulé $a(\cdot)$ ssi $\{N_0(x), x \geq 0\}$ est un processus de Poisson standard.

Preuve. Les axiomes (a) et (b) pour un processus de Poisson sont satisfaites pour N ssi elles le sont pour N_0 , car $a(\cdot)$ est continue et non-décroissante.

Reste à vérifier que $\mathbb{E}[N(t)] = a(t)$ pour $t > 0$ ssi $\mathbb{E}[N_0(x)] = x$ pour $x > 0$.

Standardisation

Un processus non-stationnaire peut être transformé en processus stationnaire simplement en étirant et comprimant l'échelle du temps.

Supposons que l'on veut un processus de Poisson de taux $\lambda(t)$ et de taux cumulé $a(t)$ au temps t , pour tout $t \geq 0$.

Soit $N_0 = \{N_0(x), x \geq 0\}$ un processus de Poisson standard (de taux 1) et définissons le processus $N = \{N(t), t \geq 0\}$ via $N(t) = N_0(a(t))$ pour $t \geq 0$. Si les sauts de N sont aux temps T_1, T_2, \dots et ceux de N_0 sont à X_1, X_2, \dots , alors on a $X_j = a(T_j)$, ou encore $T_j = a^{-1}(X_j)$.

Proposition $\{N(t), t \geq 0\}$ est un processus de Poisson de taux cumulé $a(\cdot)$ ssi $\{N_0(x), x \geq 0\}$ est un processus de Poisson standard.

Preuve. Les axiomes (a) et (b) pour un processus de Poisson sont satisfaites pour N ssi elles le sont pour N_0 , car $a(\cdot)$ est continue et non-décroissante.

Reste à vérifier que $\mathbb{E}[N(t)] = a(t)$ pour $t > 0$ ssi $\mathbb{E}[N_0(x)] = x$ pour $x > 0$.

Si $\mathbb{E}[N_0(x)] = x$, alors $\mathbb{E}[N(t)] = \mathbb{E}[N_0(a(t))] = a(t)$. Inversement, en posant $a(t) = x$ on a $N_0(x) = N(a^{-1}(x))$. Donc si $\mathbb{E}[N(t)] = a(t)$ alors $\mathbb{E}[N_0(x)] = \mathbb{E}[N(a^{-1}(x))] = a(a^{-1}(x)) = x$.

Pour **simuler** un processus de Poisson quelconque, si on sait calculer a^{-1} : on génère les X_j , puis on calcule les $T_j = a^{-1}(X_j)$.

Superposition de processus de Poisson: Si $N_1(\cdot), \dots, N_k(\cdot)$ sont des processus de Poisson indépendants de taux $\lambda_1(\cdot), \dots, \lambda_k(\cdot)$, alors $N(\cdot) = N_1(\cdot) + \dots + N_k(\cdot)$ est un processus de Poisson de taux $\lambda(\cdot) = \lambda_1(\cdot) + \dots + \lambda_k(\cdot)$.

Superposition de processus de Poisson: Si $N_1(\cdot), \dots, N_k(\cdot)$ sont des processus de Poisson indépendants de taux $\lambda_1(\cdot), \dots, \lambda_k(\cdot)$, alors $N(\cdot) = N_1(\cdot) + \dots + N_k(\cdot)$ est un processus de Poisson de taux $\lambda(\cdot) = \lambda_1(\cdot) + \dots + \lambda_k(\cdot)$.

Décomposition: inversement, si $N(\cdot)$ est un processus de Poisson de taux $\lambda(\cdot)$ et si chaque arrivée est de type j avec probabilité p_j , (indépendamment du temps et du passé), et $N_j(t)$ est le nombre d'arrivées de type j durant $(0, t]$, alors $N_j(\cdot)$ est un processus de Poisson de taux $p_j\lambda(\cdot)$.

Superposition de processus de Poisson: Si $N_1(\cdot), \dots, N_k(\cdot)$ sont des processus de Poisson indépendants de taux $\lambda_1(\cdot), \dots, \lambda_k(\cdot)$, alors $N(\cdot) = N_1(\cdot) + \dots + N_k(\cdot)$ est un processus de Poisson de taux $\lambda(\cdot) = \lambda_1(\cdot) + \dots + \lambda_k(\cdot)$.

Décomposition: inversement, si $N(\cdot)$ est un processus de Poisson de taux $\lambda(\cdot)$ et si chaque arrivée est de type j avec probabilité p_j , (indépendamment du temps et du passé), et $N_j(t)$ est le nombre d'arrivées de type j durant $(0, t]$, alors $N_j(\cdot)$ est un processus de Poisson de taux $p_j \lambda(\cdot)$.

Processus de Poisson **composé**: À chaque instant d'arrivée T_i , $N(\cdot)$ augmente (saute) de D_i , où les D_i sont i.i.d. et indépendants des T_i (arrivées en groupes).

Si $\{\lambda(t), t \geq 0\}$ est lui même un processus stochastique, $\{N(t), t \geq 0\}$ devient un processus de Poisson **doublement stochastique**, ou processus de **Cox**.

Dans ce cas, $\{N(t), t \geq 0\}$ lui-même n'est un processus de Poisson que **conditionnellement** à $\{\lambda(t), t \geq 0\}$.

Exemples: centre d'appel, comptoir de glaces, etc.

Mouvement Brownien

Un mouvement Brownien (MB) en d dimensions, avec vecteur de dérive μ et matrice de covariance Σ , est un processus

$\mathbf{X} = \{\mathbf{X}(t) = (X_1(t), \dots, X_d(t)) \in \mathbb{R}^d, t \geq 0\}$ tel que:

- (a) $\mathbf{X}(0) = 0$;
- (b) si $s \geq 0$ et $t > 0$, alors $\mathbf{X}(s+t) - \mathbf{X}(s) \sim N(\mu t, \Sigma t)$;
- (c) pour des intervalles disjoints $(t_1, t_2], \dots, (t_{2k-1}, t_{2k}]$,
les incréments $\mathbf{X}(t_2) - \mathbf{X}(t_1), \dots, \mathbf{X}(t_{2k}) - \mathbf{X}(t_{2k-1})$ sont indépendants.

Mouvement Brownien

Un mouvement Brownien (MB) en d dimensions, avec vecteur de dérive μ et matrice de covariance Σ , est un processus

$\mathbf{X} = \{\mathbf{X}(t) = (X_1(t), \dots, X_d(t)) \in \mathbb{R}^d, t \geq 0\}$ tel que:

- (a) $\mathbf{X}(0) = \mathbf{0}$;
- (b) si $s \geq 0$ et $t > 0$, alors $\mathbf{X}(s+t) - \mathbf{X}(s) \sim N(\mu t, \Sigma t)$;
- (c) pour des intervalles disjoints $(t_1, t_2], \dots, (t_{2k-1}, t_{2k}]$, les incréments $\mathbf{X}(t_2) - \mathbf{X}(t_1), \dots, \mathbf{X}(t_{2k}) - \mathbf{X}(t_{2k-1})$ sont indépendants.

Si $\mu = \mathbf{0}$, $\Sigma = \mathbf{I}$ et $\mathbf{X}(0) = \mathbf{0}$, on a un MB multivarié standard. C'est simplement un vecteur de MB standards indépendants.

Si on décompose $\Sigma = \mathbf{A}\mathbf{A}^\dagger$, alors

$$\mathbf{X}(t) = \mathbf{X}(0) + \mu t + \mathbf{A} \cdot \mathbf{B}(t)$$

où $\{\mathbf{B}(t), t \geq 0\}$ est un MB standard en d dim.

Lorsque $d = 1$, on note $\sigma_{11}^2 = \sigma^2$ (le coefficient de diffusion).

Générer un squelette du processus:

Pour générer une trajectoire observée aux instants $0 = t_0 < t_1 < \dots < t_c$:

$$\mathbf{B}(t_j) - \mathbf{B}(t_{j-1}) \sim N(\mathbf{0}, (t_j - t_{j-1})\mathbf{I}),$$

pour $j = 1, \dots, c$, puis calculer les $\mathbf{X}(t_j)$.

C'est une **marche aléatoire** multivariée.

Générer un squelette du processus:

Pour générer une trajectoire observée aux instants $0 = t_0 < t_1 < \dots < t_c$:

$$\mathbf{B}(t_j) - \mathbf{B}(t_{j-1}) \sim N(\mathbf{0}, (t_j - t_{j-1})\mathbf{I}),$$

pour $j = 1, \dots, c$, puis calculer les $\mathbf{X}(t_j)$.

C'est une **marche aléatoire** multivariée.

Équivalent: générer

$$Z_{1,1}, \dots, Z_{1,d}, \dots, Z_{c,1}, \dots, Z_{c,d} \quad \text{i.i.d. } N(0, 1)$$

et poser

$$\mathbf{X}(t_j) = \mathbf{X}(t_{j-1}) + (t_j - t_{j-1})\boldsymbol{\mu} + \sqrt{t_j - t_{j-1}}\mathbf{A}(Z_{j,1}, \dots, Z_{j,d})^t$$

pour $j = 1, \dots, c$.

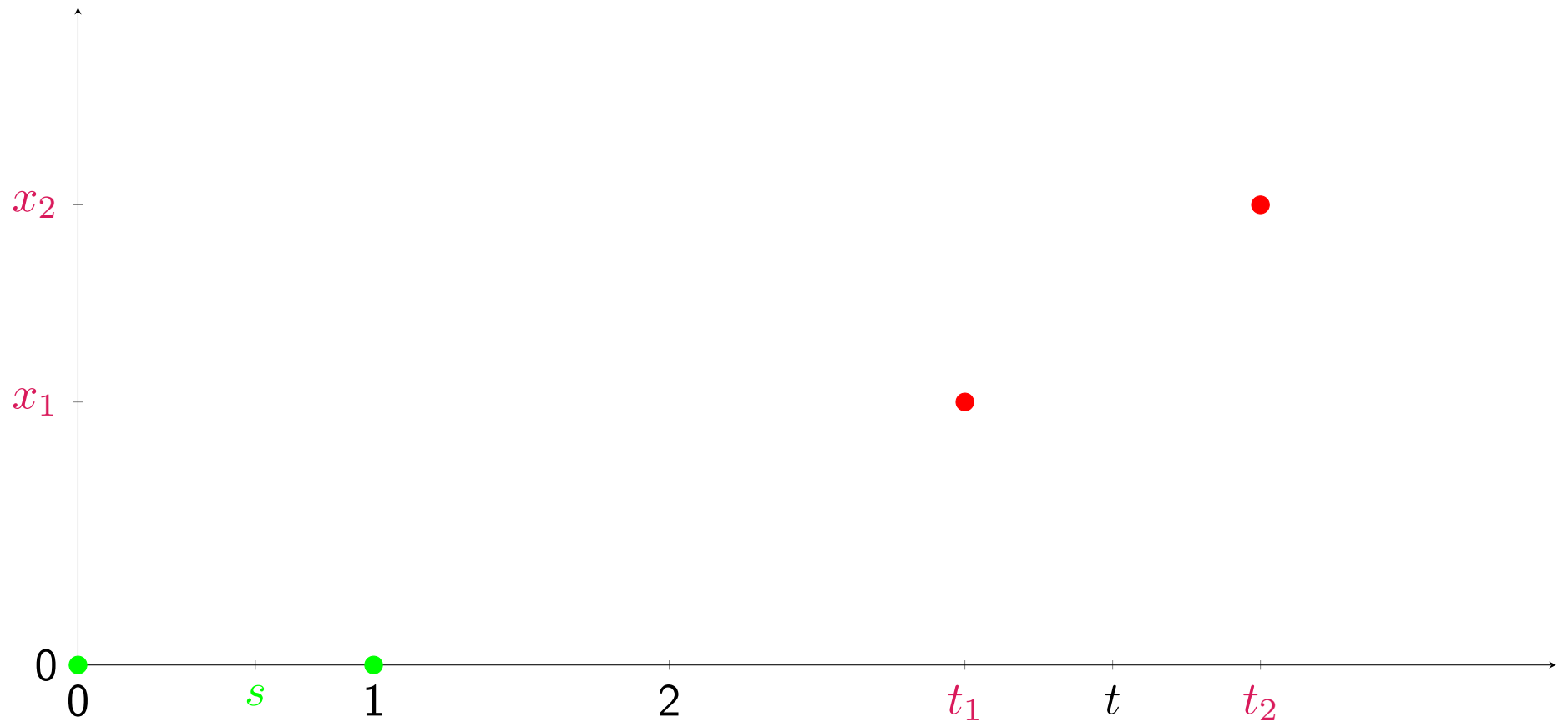
En général, on a $\mathbb{E}[X_i(t)] = t\mu_i$ et $\text{Cov}[X_i(s), X_j(t)] = \min(s, t)\sigma_{i,j}$.

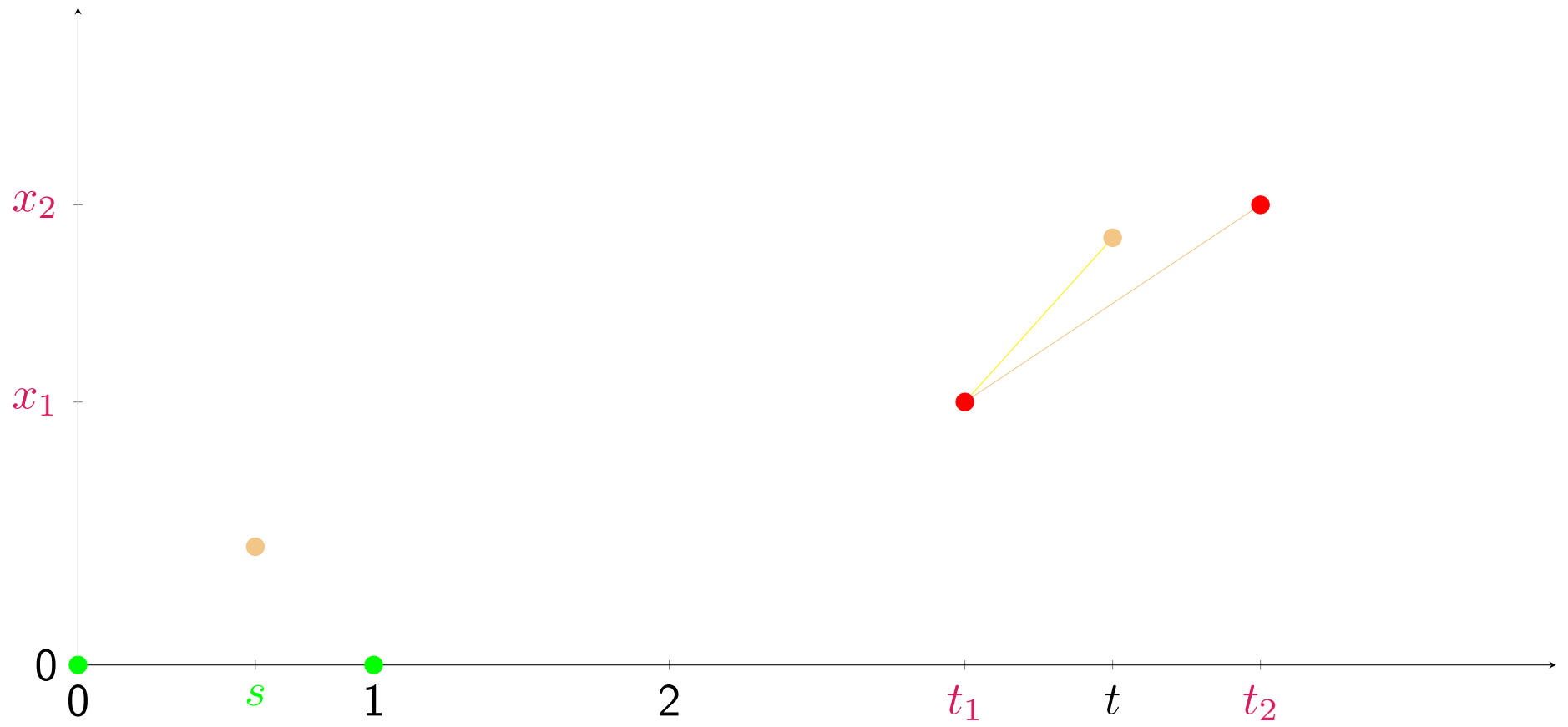
Ainsi, pour des instants d'observation $0 = t_0 < t_1 < \dots < t_c$ quelconques, le vecteur $\mathbf{Y} = (X_1(t_1), \dots, X_d(t_1), \dots, X_1(t_c), \dots, X_d(t_c))^t$ est multinormal de moyenne $\boldsymbol{\mu}_y$ et de covariance $\boldsymbol{\Sigma}_y$ dont les éléments sont connus.

Si on décompose $\boldsymbol{\Sigma}_y = \mathbf{A}_y \mathbf{A}_y^t$ (comme on veut), on peut générer \mathbf{Y} via

$$\mathbf{Y} = \boldsymbol{\mu}_y + \mathbf{A}_y \mathbf{Z},$$

où $\mathbf{Z} \sim \text{Normale}(\mathbf{0}, \mathbf{I})$ en cd dimensions.





Pont Brownien

Si on conditionne sur $\{\mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2\}$, le processus $\mathbf{X}^0 = \{X^0(t), t_1 \leq t \leq t_2\}$ est un **pont brownien** (multivarié).

Pont Brownien

Si on conditionne sur $\{\mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2\}$, le processus $\mathbf{X}^0 = \{X^0(t), t_1 \leq t \leq t_2\}$ est un **pont brownien** (multivarié).

Dans le cas où X est un MB standard, $t_1 = 0$, $t_2 = 1$, et $\mathbf{X}(0) = \mathbf{X}(1) = 0$, on a un **pont Brownien standard**.

Pont Brownien

Si on conditionne sur $\{\mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2\}$, le processus $\mathbf{X}^0 = \{X^0(t), t_1 \leq t \leq t_2\}$ est un **pont brownien** (multivarié).

Dans le cas où X est un MB standard, $t_1 = 0$, $t_2 = 1$, et $\mathbf{X}(0) = \mathbf{X}(1) = 0$, on a un **pont Brownien standard**.

Proposition: Pour $t_1 < t < t_2$, conditionnellement à $\mathbf{X}(t_1) = \mathbf{x}_1$ et $\mathbf{X}(t_2) = \mathbf{x}_2$, $\mathbf{X}(t)$ est multinormal avec

$$\begin{aligned}\mathbb{E}[\mathbf{X}(t) \mid \mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2] &= \mathbf{x}_1 + \frac{t - t_1}{t_2 - t_1}(\mathbf{x}_2 - \mathbf{x}_1), \\ \text{Cov}[\mathbf{X}(t) \mid \mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2] &= \frac{(t - t_1)(t_2 - t)}{t_2 - t_1} \boldsymbol{\Sigma}.\end{aligned}$$

Pont Brownien

Si on conditionne sur $\{\mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2\}$, le processus $\mathbf{X}^0 = \{X^0(t), t_1 \leq t \leq t_2\}$ est un **pont brownien** (multivarié).

Dans le cas où X est un MB standard, $t_1 = 0$, $t_2 = 1$, et $\mathbf{X}(0) = \mathbf{X}(1) = 0$, on a un **pont Brownien standard**.

Proposition: Pour $t_1 < t < t_2$, conditionnellement à $\mathbf{X}(t_1) = \mathbf{x}_1$ et $\mathbf{X}(t_2) = \mathbf{x}_2$, $\mathbf{X}(t)$ est multinormal avec

$$\begin{aligned}\mathbb{E}[\mathbf{X}(t) \mid \mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2] &= \mathbf{x}_1 + \frac{t - t_1}{t_2 - t_1}(\mathbf{x}_2 - \mathbf{x}_1), \\ \text{Cov}[\mathbf{X}(t) \mid \mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2] &= \frac{(t - t_1)(t_2 - t)}{t_2 - t_1} \boldsymbol{\Sigma}.\end{aligned}$$

Simulation par **raffinements successifs** (supposons que c est une puissance de 2):
générer $\mathbf{X}(t_c) \sim N(t_c \boldsymbol{\mu}, t_c \boldsymbol{\Sigma})$;
puis générer $\mathbf{X}(t_{c/2})$, selon sa loi conditionnelle à $(\mathbf{X}(0), \mathbf{X}(t_c))$, une multinormale de moyenne $\mathbf{X}(0) + (\mathbf{X}(t_c) - \mathbf{X}(0))t_{c/2}/t_c$ et de covariance $(t_{c/2}(t_c - t_{c/2})/t_c)\boldsymbol{\Sigma}$;

Pont Brownien

Si on conditionne sur $\{\mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2\}$, le processus $\mathbf{X}^0 = \{X^0(t), t_1 \leq t \leq t_2\}$ est un **pont brownien** (multivarié).

Dans le cas où X est un MB standard, $t_1 = 0$, $t_2 = 1$, et $\mathbf{X}(0) = \mathbf{X}(1) = 0$, on a un **pont Brownien standard**.

Proposition: Pour $t_1 < t < t_2$, conditionnellement à $\mathbf{X}(t_1) = \mathbf{x}_1$ et $\mathbf{X}(t_2) = \mathbf{x}_2$, $\mathbf{X}(t)$ est multinormal avec

$$\begin{aligned}\mathbb{E}[\mathbf{X}(t) \mid \mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2] &= \mathbf{x}_1 + \frac{t - t_1}{t_2 - t_1}(\mathbf{x}_2 - \mathbf{x}_1), \\ \text{Cov}[\mathbf{X}(t) \mid \mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{x}_2] &= \frac{(t - t_1)(t_2 - t)}{t_2 - t_1} \Sigma.\end{aligned}$$

Simulation par **raffinements successifs** (supposons que c est une puissance de 2):
générer $\mathbf{X}(t_c) \sim N(t_c \boldsymbol{\mu}, t_c \boldsymbol{\Sigma})$;
puis générer $\mathbf{X}(t_{c/2})$, selon sa loi conditionnelle à $(\mathbf{X}(0), \mathbf{X}(t_c))$, une multinormale de moyenne $\mathbf{X}(0) + (\mathbf{X}(t_c) - \mathbf{X}(0))t_{c/2}/t_c$ et de covariance $(t_{c/2}(t_c - t_{c/2})/t_c)\boldsymbol{\Sigma}$;
continuer récursivement pour générer $\mathbf{X}(t_{c/4})$ conditionnellement à $(\mathbf{X}(0), \mathbf{X}(t_{c/2}))$, puis $\mathbf{X}(t_{3c/4})$ conditionnellement à $(\mathbf{X}(t_{c/2}), \mathbf{X}(t_c))$, etc.

Mouvement Brownien géométrique

$\{\mathbf{S}(t) = (S_1(t), \dots, S_d(t)), t \geq 0\}$ est un mouvement brownien géométrique (MBG) (multivarié) de paramètres $\boldsymbol{\mu}$ et $\boldsymbol{\Sigma}$ si $\{\mathbf{X}(t) = \ln \mathbf{S}(t) - \ln(\mathbf{S}(0)), t \geq 0\}$ est un MB multivarié de dérive $\boldsymbol{\mu} - (\sigma_1^2/2, \dots, \sigma_d^2/2)^t$ et covariance $\boldsymbol{\Sigma}$. On peut alors écrire

$$S_i(t) = S_i(0) \exp[X_i(t)]$$

pour $i = 1, \dots, d$. Pour générer $\mathbf{S}(\cdot)$, il suffit de générer $\mathbf{X}(\cdot)$.

Ce processus ne prend jamais de valeurs négatives.

Utilisé dans le modèle de Black-Scholes en finance et économie.

Mouvement Brownien géométrique

$\{\mathbf{S}(t) = (S_1(t), \dots, S_d(t)), t \geq 0\}$ est un mouvement brownien géométrique (MBG) (multivarié) de paramètres $\boldsymbol{\mu}$ et $\boldsymbol{\Sigma}$ si $\{\mathbf{X}(t) = \ln \mathbf{S}(t) - \ln(\mathbf{S}(0)), t \geq 0\}$ est un MB multivarié de dérive $\boldsymbol{\mu} - (\sigma_1^2/2, \dots, \sigma_d^2/2)^t$ et covariance $\boldsymbol{\Sigma}$. On peut alors écrire

$$S_i(t) = S_i(0) \exp[X_i(t)]$$

pour $i = 1, \dots, d$. Pour générer $\mathbf{S}(\cdot)$, il suffit de générer $\mathbf{X}(\cdot)$.

Ce processus ne prend jamais de valeurs négatives.

Utilisé dans le modèle de Black-Scholes en finance et économie.

En une dimension, cela donne

$$S(t) = S(0) \exp[X(t)] = S(0) \exp[(\mu - \sigma^2/2)t + \sigma B(t)],$$

où $\mu = \mu_1$ est la dérive et $\sigma^2 = \sigma_{1,1}$ la volatilité.

On a $S(t_0 + t)/S(t_0) \sim \text{Lognormale}((\mu - \sigma^2/2)t, \sigma^2 t)$.

Le même processus peut aussi être défini via

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dB(t).$$

Processus Gaussien

Généralisations du MB: Le vecteur $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}(t)]/t$ et la matrice de covariance $\text{Cov}[\mathbf{X}(t)]/t$ peuvent dépendre de t .

Processus Gaussien

Généralisations du MB: Le vecteur $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}(t)]/t$ et la matrice de covariance $\text{Cov}[\mathbf{X}(t)]/t$ peuvent dépendre de t .

Plus général: On remplace le temps $t \in [0, \infty)$ par un indice arbitraire $t \in \mathcal{I}$, qui peut indiquer une position dans l'espace, ou une combinaison de temps et d'espace, etc.

Le processus $\{X(t), t \in \mathcal{I}\}$ est un **processus Gaussien** si quelque soient $t_1, \dots, t_c \in \mathcal{I}$, la loi conjointe de $\mathbf{X}_c = (X(t_1), \dots, X(t_c))^t$ est multinormale.

Processus Gaussien

Généralisations du MB: Le vecteur $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}(t)]/t$ et la matrice de covariance $\text{Cov}[\mathbf{X}(t)]/t$ peuvent dépendre de t .

Plus général: On remplace le temps $t \in [0, \infty)$ par un indice arbitraire $t \in \mathcal{I}$, qui peut indiquer une position dans l'espace, ou une combinaison de temps et d'espace, etc.

Le processus $\{X(t), t \in \mathcal{I}\}$ est un **processus Gaussien** si quelque soient $t_1, \dots, t_c \in \mathcal{I}$, la loi conjointe de $\mathbf{X}_c = (X(t_1), \dots, X(t_c))^t$ est multinormale.

On peut le spécifier par sa fonction de moyenne $\{\mu(t), t \in \mathcal{I}\}$ et sa fonction de covariance $\text{Cov}[X(s), X(t)]$, pour $s, t \in \mathcal{I}$, qui doit être définie non négative: $\mathbf{x} \text{Cov}(\mathbf{X}_c) \mathbf{x}^t \geq 0$ pour tout $\mathbf{x} \in \mathbb{R}^c$.

Processus Gaussien

Généralisations du MB: Le vecteur $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}(t)]/t$ et la matrice de covariance $\text{Cov}[\mathbf{X}(t)]/t$ peuvent dépendre de t .

Plus général: On remplace le temps $t \in [0, \infty)$ par un indice arbitraire $t \in \mathcal{I}$, qui peut indiquer une position dans l'espace, ou une combinaison de temps et d'espace, etc.

Le processus $\{X(t), t \in \mathcal{I}\}$ est un **processus Gaussien** si quelque soient $t_1, \dots, t_c \in \mathcal{I}$, la loi conjointe de $\mathbf{X}_c = (X(t_1), \dots, X(t_c))^t$ est multinormale.

On peut le spécifier par sa fonction de moyenne $\{\mu(t), t \in \mathcal{I}\}$ et sa fonction de covariance $\text{Cov}[X(s), X(t)]$, pour $s, t \in \mathcal{I}$, qui doit être définie non négative: $\mathbf{x} \text{Cov}(\mathbf{X}_c) \mathbf{x}^t \geq 0$ pour tout $\mathbf{x} \in \mathbb{R}^c$.

Si $\mathbb{E}[\mathbf{X}_c]$ et $\boldsymbol{\Sigma}_c = \text{Cov}[\mathbf{X}_c]$ sont disponibles explicitement, on peut générer le vecteur multinormal \mathbf{X}_c selon les techniques habituelles:

On décompose $\boldsymbol{\Sigma}_c = \mathbf{A}_c \mathbf{A}_c^t$ par notre méthode favorite, on génère un vecteur \mathbf{Z}_c de c normales standard indépendantes, et on pose $\mathbf{X}_c = (\mu(t_1), \dots, \mu(t_c))^t + \mathbf{A}_c \mathbf{Z}_c$.

Équation différentielle stochastique (EDS)

On considère $\mathbf{X} = \{\mathbf{X}(t), t \geq 0\}$ dans \mathbb{R}^d défini via l'EDS

$$d\mathbf{X}(t) = \boldsymbol{\mu}(\mathbf{X}(t), t) \cdot dt + \mathbf{A}(\mathbf{X}(t), t) \cdot d\mathbf{B}(t),$$

avec $\mathbf{X}(0) = \mathbf{x}_0 \in \mathbb{R}^d$, où \mathbf{B} est un MB standard en q dimensions,

$\boldsymbol{\mu}(\mathbf{X}(t), t) \in \mathbb{R}^d$ (la dérive), et $\mathbf{A}(\mathbf{X}(t), t)$ est une matrice $d \times q$.

Coefficient de diffusion (covariance): $\boldsymbol{\Sigma}(\mathbf{X}(t), t) = \mathbf{A}(\mathbf{X}(t), t)\mathbf{A}(\mathbf{X}(t), t)^t$.

Équation différentielle stochastique (EDS)

On considère $\mathbf{X} = \{\mathbf{X}(t), t \geq 0\}$ dans \mathbb{R}^d défini via l'EDS

$$d\mathbf{X}(t) = \boldsymbol{\mu}(\mathbf{X}(t), t) \cdot dt + \mathbf{A}(\mathbf{X}(t), t) \cdot d\mathbf{B}(t),$$

avec $\mathbf{X}(0) = \mathbf{x}_0 \in \mathbb{R}^d$, où \mathbf{B} est un MB standard en q dimensions,

$\boldsymbol{\mu}(\mathbf{X}(t), t) \in \mathbb{R}^d$ (la dérive), et $\mathbf{A}(\mathbf{X}(t), t)$ est une matrice $d \times q$.

Coefficient de diffusion (covariance): $\Sigma(\mathbf{X}(t), t) = \mathbf{A}(\mathbf{X}(t), t)\mathbf{A}(\mathbf{X}(t), t)^t$.

Définition équivalente (équation d'Itô):

$$\mathbf{X}(t) = \mathbf{X}(0) + \int_0^t \boldsymbol{\mu}(\mathbf{X}(s), s)ds + \int_0^t \mathbf{A}(\mathbf{X}(s), s)d\mathbf{B}(s).$$

Cas homogène: $\boldsymbol{\mu}$ et \mathbf{A} ne dépendent pas de t .

Example. Si μ et \mathbf{A} sont constants, on obtient un MB ordinaire:

$$d\mathbf{X}(t) = \mu dt + \mathbf{A}dB(t) \quad \text{et} \quad \mathbf{X}(t) = \mathbf{X}(0) + \mu t + \mathbf{A}\mathbf{B}(t).$$

On obtient un MBG (1-dim) via

$$dX(t) = \mu X(t)dt + \sigma X(t)dB(t),$$

ou encore

$$X(t) = X(0) \exp [(\mu - \sigma^2/2)t + \sigma B(t)] .$$

Example. Si μ et \mathbf{A} sont constants, on obtient un MB ordinaire:

$$d\mathbf{X}(t) = \mu dt + \mathbf{A}dB(t) \quad \text{et} \quad \mathbf{X}(t) = \mathbf{X}(0) + \mu t + \mathbf{A}\mathbf{B}(t).$$

On obtient un MBG (1-dim) via

$$dX(t) = \mu X(t)dt + \sigma X(t)dB(t),$$

ou encore

$$X(t) = X(0) \exp [(\mu - \sigma^2/2)t + \sigma B(t)] .$$

Si on connaît la loi exacte de $\mathbf{X}(t+s)$ sachant $\mathbf{X}(t)$, alors on peut simuler le processus à des instants $0 = t_0 < t_1 < t_2 < \dots$. C'est le cas pour le MB, MBG, et les processus O-U et CIR, par exemple.

Mais en général on doit approximer en discrétisant le temps.

Méthode d'Euler: On discrétise le temps en pas de longueur $dt = h$. On a $\tilde{\mathbf{X}}(0) = \mathbf{X}(0)$ et

$$\tilde{\mathbf{X}}((j+1)h) = \tilde{\mathbf{X}}(jh) + \boldsymbol{\mu}(\tilde{\mathbf{X}}(jh), jh) h + \mathbf{A}(\tilde{\mathbf{X}}(jh), jh) \sqrt{h} \mathbf{Z}_j,$$

pour $j = 1, 2, \dots$, où les \mathbf{Z}_j sont i.i.d. Normale($\mathbf{0}, \mathbf{I}$).

Méthode d'Euler: On discrétise le temps en pas de longueur $dt = h$. On a $\tilde{\mathbf{X}}(0) = \mathbf{X}(0)$ et

$$\tilde{\mathbf{X}}((j+1)h) = \tilde{\mathbf{X}}(jh) + \boldsymbol{\mu}(\tilde{\mathbf{X}}(jh), jh) h + \mathbf{A}(\tilde{\mathbf{X}}(jh), jh) \sqrt{h} \mathbf{Z}_j,$$

pour $j = 1, 2, \dots$, où les \mathbf{Z}_j sont i.i.d. Normale($\mathbf{0}, \mathbf{I}$).

Méthode de Milstein: Ajoute une correction pour tenir compte du fait que $\mathbf{A}(\mathbf{X}(t), t)$ varie dans pour $t \in [jh, jh + h]$. Compliqué pour $d > 1$. Pour $d = 1$:

$$dX(t) = \mu(X(t), t) dt + \sigma(X(t), t) dB(t),$$

où $\sigma = \mathbf{A}$. Avec la correction:

$$\begin{aligned} \tilde{X}((j+1)h) &= \tilde{X}(jh) + \mu(\tilde{X}(jh), jh) h + \sigma(\tilde{X}(jh), jh) \sqrt{h} Z_j \\ &\quad + (\partial\sigma/\partial x)(\tilde{X}(jh), jh) \sigma(\tilde{X}(jh), jh) (Z_j^2 - 1)h/2. \end{aligned}$$

Le gain en précision est souvent négligeable ou faible.

Processus de retour à la moyenne d'Ornstein-Uhlenbeck (O-U)

27

Processus $\{X(t), t \geq 0\}$ défini via

$$dX(t) = \alpha(b - X(t))dt + \sigma dB(t),$$

où B est MB standard, et α , b , et σ sont des constantes positives.

Ce processus est constamment attiré vers sa moyenne globale b .

C'est le [modèle de Vasicek](#) pour la variation des taux d'intérêt à court terme.

Processus de retour à la moyenne d'Ornstein-Uhlenbeck (O-U)

27

Processus $\{X(t), t \geq 0\}$ défini via

$$dX(t) = \alpha(b - X(t))dt + \sigma dB(t),$$

où B est MB standard, et α , b , et σ sont des constantes positives.

Ce processus est constamment attiré vers sa moyenne globale b .

C'est le **modèle de Vasicek** pour la variation des taux d'intérêt à court terme.

Pour $0 < s < t$, conditionnellement à $X(s) = x$, $X(t)$ est normal de moyenne $e^{-\alpha(t-s)}x + (1 - e^{-\alpha(t-s)})b$ et variance $(1 - e^{-2\alpha(t-s)})\sigma^2/(2\alpha)$.

Permet de simuler facilement le processus aux instants $0 = t_0 < t_1 < \dots < t_c$.

On a ici un processus Gaussien.

Version en d dimensions: voir les notes.

Modèle de Cox, Ingersoll, et Ross (CIR).

Semblable à O-U, mais avec un facteur $\sqrt{X(t)}$ pour la diffusion (ce qui, en particulier, empêche les valeurs négatives):

$$dX(t) = \alpha(b - X(t))dt + \sigma\sqrt{X(t)}dB(t),$$

où B est MB standard, et α , b , et σ sont des constantes positives.

Proposé par Cox, Ingersoll, et Ross (CIR) pour les taux d'intérêt.

Modèle de Cox, Ingersoll, et Ross (CIR).

Semblable à O-U, mais avec un facteur $\sqrt{X(t)}$ pour la diffusion (ce qui, en particulier, empêche les valeurs négatives):

$$dX(t) = \alpha(b - X(t))dt + \sigma\sqrt{X(t)}dB(t),$$

où B est MB standard, et α , b , et σ sont des constantes positives.

Proposé par Cox, Ingersoll, et Ross (CIR) pour les taux d'intérêt.

Pour $0 < s < t$, conditionnellement à $X(s) = x$, $X(t)$ a la même distribution que

$$\frac{\sigma^2(1 - e^{-\alpha(t-s)})}{4\alpha} Y$$

où Y est une chi-deux non centrée à $k = 4b\alpha/\sigma^2$ degrés de liberté et paramètre de non-centralité

$$\lambda = \frac{4\alpha e^{-\alpha(t-s)}x}{\sigma^2(1 - e^{-\alpha(t-s)})}.$$

Ceci permet de simuler “facilement” le processus.

Processus de Lévy

$Y = \{Y(t), t \geq 0\}$ est un processus de Lévy si ses incréments sont stationnaires et indépendants. Autrement dit, pour des intervalles disjoints $(t_{2j-1}, t_{2j}]$, $j = 1, 2, \dots$, les v.a. $X_j = Y(t_{2j}) - Y(t_{2j-1})$ sont indépendantes et la loi de X_j ne dépend que de $t_{2j} - t_{2j-1}$.

Processus de Lévy

$Y = \{Y(t), t \geq 0\}$ est un processus de Lévy si ses incréments sont stationnaires et indépendants. Autrement dit, pour des intervalles disjoints $(t_{2j-1}, t_{2j}]$, $j = 1, 2, \dots$, les v.a. $X_j = Y(t_{2j}) - Y(t_{2j-1})$ sont indépendantes et la loi de X_j ne dépend que de $t_{2j} - t_{2j-1}$.

Définition équivalente: le processus est infiniment divisible:

Pour tout t et tout n , $Y(t) - Y(0)$ s'écrit comme une somme de n v.a. i.i.d.

Exemples: processus de Poisson, brownien, gamma, etc.

Processus de Lévy

$Y = \{Y(t), t \geq 0\}$ est un processus de Lévy si ses incréments sont stationnaires et indépendants. Autrement dit, pour des intervalles disjoints $(t_{2j-1}, t_{2j}]$, $j = 1, 2, \dots$, les v.a. $X_j = Y(t_{2j}) - Y(t_{2j-1})$ sont indépendantes et la loi de X_j ne dépend que de $t_{2j} - t_{2j-1}$.

Définition équivalente: le processus est infiniment divisible:

Pour tout t et tout n , $Y(t) - Y(0)$ s'écrit comme une somme de n v.a. i.i.d.

Exemples: processus de Poisson, brownien, gamma, etc.

Tout processus de Lévy peut s'écrire comme la somme d'un Brownien et d'un processus de saut, avec hauteurs de sauts aléatoires.

Si le taux de saut λ (nombre moyen par unité de temps) est fini, on peut écrire

$$Y(t) = \mu t + \sigma B(t) + \sum_{j=1}^{N(t)} D_j \quad \text{for } t \geq 0,$$

où B est un mouvement Brownien standard, N est un processus de Poisson de taux λ , et les D_j sont des v.a. i.i.d., indépendantes de B et N .

Facile à simuler si on sait générer les D_j .

Mais dans plusieurs cas on a $\lambda = \infty$ (et la plupart des sauts sont minuscules).
Voir Asmussen et Glynn (2007).

Mais dans plusieurs cas on a $\lambda = \infty$ (et la plupart des sauts sont minuscules). Voir Asmussen et Glynn (2007).

Dans les cas examinés ici, on suppose que l'on sait générer l'incrément $Y(t)$ pour tout t . On peut alors générer le processus aux instants d'observation $0 = t_0 < t_1 < \dots < t_c$ par **échantillonnage séquentiel** (ou technique de **marche aléatoire**), en générant les incréments $Y(t_j) - Y(t_{j-1})$, pour $j = 1, \dots, c$, successivement.

Mais dans plusieurs cas on a $\lambda = \infty$ (et la plupart des sauts sont minuscules). Voir Asmussen et Glynn (2007).

Dans les cas examinés ici, on suppose que l'on sait générer l'incrément $Y(t)$ pour tout t . On peut alors générer le processus aux instants d'observation $0 = t_0 < t_1 < \dots < t_c$ par **échantillonnage séquentiel** (ou technique de **marche aléatoire**), en générant les incréments $Y(t_j) - Y(t_{j-1})$, pour $j = 1, \dots, c$, successivement.

Dans certains cas, pour $t_1 < s < t_2$, on connaît la loi de $Y(s)$ conditionnelle à $(Y(t_1), Y(t_2))$.

Mais dans plusieurs cas on a $\lambda = \infty$ (et la plupart des sauts sont minuscules). Voir Asmussen et Glynn (2007).

Dans les cas examinés ici, on suppose que l'on sait générer l'incrément $Y(t)$ pour tout t . On peut alors générer le processus aux instants d'observation $0 = t_0 < t_1 < \dots < t_c$ par **échantillonnage séquentiel** (ou technique de **marche aléatoire**), en générant les incréments $Y(t_j) - Y(t_{j-1})$, pour $j = 1, \dots, c$, successivement.

Dans certains cas, pour $t_1 < s < t_2$, on connaît la loi de $Y(s)$ conditionnelle à $(Y(t_1), Y(t_2))$. On peut alors, en principe, simuler la trajectoire sur $[0, t]$ par raffinements successifs ("**Lévy bridge sampling**") : générer d'abord $Y(t)$;
 puis $Y(t/2)$ conditionnellement à $(Y(0), Y(t))$;
 puis $Y(t/4)$ conditionnellement à $(Y(0), Y(t/2))$;
 puis $Y(3t/4)$ conditionnellement à $(Y(t/2), Y(t))$;
 puis $Y(t/8)$ conditionnellement à $(Y(0), Y(t/4))$; etc.

Changement de l'échelle de temps.

On a un processus quelconque $X = \{X(t), t \geq 0\}$.

On choisit $a : [0, \infty) \rightarrow [0, \infty)$ non décroissante. L'horloge avance à la vitesse $a'(t)$ au temps t (si $a'(t)$ existe).

Le processus X est modifié en $Y = \{Y(t), t \geq 0\}$ où $Y(t) = X(a(t))$.

Changement de l'échelle de temps.

On a un processus quelconque $X = \{X(t), t \geq 0\}$.

On choisit $a : [0, \infty) \rightarrow [0, \infty)$ non décroissante. L'horloge avance à la vitesse $a'(t)$ au temps t (si $a'(t)$ existe).

Le processus X est modifié en $Y = \{Y(t), t \geq 0\}$ où $Y(t) = X(a(t))$.

Horloge aléatoire (“random time change”).

Changement aléatoire non-linéaire de l'échelle du temps.

L'horloge est un processus (subordonateur) non-décroissant $T = \{T(t), t \geq 0\}$, avec $T(0) = 0$.

On remplace $X(t)$ par $Y(t) = X(T(t))$ pour chaque t , pour obtenir $Y = \{Y(t), t \geq 0\}$.

Si X et T sont des processus de Lévy, alors Y en est un aussi.

Changement de l'échelle de temps.

On a un processus quelconque $X = \{X(t), t \geq 0\}$.

On choisit $a : [0, \infty) \rightarrow [0, \infty)$ non décroissante. L'horloge avance à la vitesse $a'(t)$ au temps t (si $a'(t)$ existe).

Le processus X est modifié en $Y = \{Y(t), t \geq 0\}$ où $Y(t) = X(a(t))$.

Horloge aléatoire (“random time change”).

Changement aléatoire non-linéaire de l'échelle du temps.

L'horloge est un processus (subordonateur) non-décroissant $T = \{T(t), t \geq 0\}$, avec $T(0) = 0$.

On remplace $X(t)$ par $Y(t) = X(T(t))$ pour chaque t , pour obtenir $Y = \{Y(t), t \geq 0\}$.

Si X et T sont des processus de Lévy, alors Y en est un aussi.

Si X est un mouvement Brownien, cela est équivalent à remplacer le paramètre de volatilité σ par un processus de volatilité stochastique $\{\sigma(t), t \geq 0\}$.

Processus gamma

Un processus gamma $G = \{G(t), t \geq 0\}$ est un processus de Lévy dont l'accroissement sur un intervalle de longueur t suit une loi gamma de paramètres $(t\alpha, \lambda) = (t\mu^2/\nu, \mu/\nu)$ (moyenne $t\mu$ et variance $t\nu$).

Assez facile à simuler par marche aléatoire.

Processus gamma

Un processus gamma $G = \{G(t), t \geq 0\}$ est un processus de Lévy dont l'accroissement sur un intervalle de longueur t suit une loi gamma de paramètres $(t\alpha, \lambda) = (t\mu^2/\nu, \mu/\nu)$ (moyenne $t\mu$ et variance $t\nu$).

Assez facile à simuler par marche aléatoire.

Pour $t_a < t < t_b$, la loi de $(G(t) - G(t_a))/(G(t_b) - G(t_a))$ conditionnelle à $(G(t_a), G(t_b))$ est beta de paramètres $((t - t_a)\alpha, (t_b - t)\alpha)$.

On peut donc aussi le simuler facilement par pont gamma (bisection).

Processus gamma

Un processus gamma $G = \{G(t), t \geq 0\}$ est un processus de Lévy dont l'accroissement sur un intervalle de longueur t suit une loi gamma de paramètres $(t\alpha, \lambda) = (t\mu^2/\nu, \mu/\nu)$ (moyenne $t\mu$ et variance $t\nu$).

Assez facile à simuler par marche aléatoire.

Pour $t_a < t < t_b$, la loi de $(G(t) - G(t_a))/(G(t_b) - G(t_a))$ conditionnelle à $(G(t_a), G(t_b))$ est beta de paramètres $((t - t_a)\alpha, (t_b - t)\alpha)$.

On peut donc aussi le simuler facilement par pont gamma (bisection).

Ce processus a une trajectoire non décroissante. On peut donc l'utiliser comme subordonateur pour changer l'échelle de temps.

Processus variance gamma

Un processus variance-gamma (VG) $Y = \{Y(t), t \geq 0\}$ est défini par

$$Y(t) = X(G(t))$$

où X est un MB de paramètres μ et σ^2 , G est un processus gamma de paramètres 1 (moyenne) et ν (variance), et X et G sont indépendants.

Processus variance gamma

Un processus variance-gamma (VG) $Y = \{Y(t), t \geq 0\}$ est défini par

$$Y(t) = X(G(t))$$

où X est un MB de paramètres μ et σ^2 , G est un processus gamma de paramètres 1 (moyenne) et ν (variance), et X et G sont indépendants.

Remplacer le MB dans un MBG par un processus VG donne un modèle plus flexible et plus réaliste pour modéliser l'évolution des prix d'actifs.

Processus variance gamma

Un processus **variance-gamma (VG)** $Y = \{Y(t), t \geq 0\}$ est défini par

$$Y(t) = X(G(t))$$

où X est un MB de paramètres μ et σ^2 , G est un processus gamma de paramètres 1 (moyenne) et ν (variance), et X et G sont indépendants.

Remplacer le MB dans un MBG par un processus VG donne un modèle plus flexible et plus réaliste pour modéliser l'évolution des prix d'actifs.

Pour **simuler** aux instants $0 = t_0 < t_1 < \dots < t_c$, on peut générer G puis X de manière séquentielle:

générer $\tau_1 = G(t_1)$, $X(\tau_1)$, $\tau_2 = G(t_2)$, $X(\tau_2)$, etc., dans cet ordre.

Processus variance gamma

Un processus **variance-gamma (VG)** $Y = \{Y(t), t \geq 0\}$ est défini par

$$Y(t) = X(G(t))$$

où X est un MB de paramètres μ et σ^2 , G est un processus gamma de paramètres 1 (moyenne) et ν (variance), et X et G sont indépendants.

Remplacer le MB dans un MBG par un processus VG donne un modèle plus flexible et plus réaliste pour modéliser l'évolution des prix d'actifs.

Pour **simuler** aux instants $0 = t_0 < t_1 < \dots < t_c$, on peut générer G puis X de manière séquentielle:

générer $\tau_1 = G(t_1)$, $X(\tau_1)$, $\tau_2 = G(t_2)$, $X(\tau_2)$, etc., dans cet ordre.

Autre stratégie: raffinements successifs. Générer, dans cet ordre:

$$\begin{aligned} \tau_c &= G(t_c), X(\tau_c), \tau_{c/2} = G(t_{c/2}), X(\tau_{c/2}), \\ \tau_{c/4} &= G(t_{c/4}), X(\tau_{c/4}), \tau_{3c/4} = G(t_{3c/4}), X(\tau_{3c/4}), \dots, \end{aligned}$$

Un processus VG Y peut aussi s'écrire comme

$$Y(t) = G^+(t) - G^-(t),$$

où G^+ et G^- sont des processus gamma indépendants de paramètres (μ^+, ν^+) et (μ^-, ν^-) , avec

$$\mu^+ = (\sqrt{\theta^2 + 2\sigma^2/\nu} + \theta)/2,$$

$$\mu^- = (\sqrt{\theta^2 + 2\sigma^2/\nu} - \theta)/2,$$

$$\nu^+ = (\mu^+)^2\nu, \text{ and}$$

$$\nu^- = (\mu^-)^2\nu.$$

On peut donc simuler Y en simulant G^+ et G^- , soit l'un après l'autre, ou encore comme ceci: générer, dans cet ordre,

$G^+(t_c)$, $G^-(t_c)$, $G^+(t_{c/2})$, $G^-(t_{c/2})$, $G^+(t_{c/4})$, $G^-(t_{c/4})$, etc.

Un processus VG Y peut aussi s'écrire comme

$$Y(t) = G^+(t) - G^-(t),$$

où G^+ et G^- sont des processus gamma indépendants de paramètres (μ^+, ν^+) et (μ^-, ν^-) , avec

$$\mu^+ = (\sqrt{\theta^2 + 2\sigma^2/\nu} + \theta)/2,$$

$$\mu^- = (\sqrt{\theta^2 + 2\sigma^2/\nu} - \theta)/2,$$

$$\nu^+ = (\mu^+)^2\nu, \text{ and}$$

$$\nu^- = (\mu^-)^2\nu.$$

On peut donc simuler Y en simulant G^+ et G^- , soit l'un après l'autre, ou encore comme ceci: générer, dans cet ordre,

$G^+(t_c)$, $G^-(t_c)$, $G^+(t_{c/2})$, $G^-(t_{c/2})$, $G^+(t_{c/4})$, $G^-(t_{c/4})$, etc.

Le fait que G^+ et G^- sont non décroissants nous donne des bornes sur $Y(t) = G^+(t) - G^-(t)$ pour tout $t \leq t_c$, à chaque étape de cette simulation.

Processus inverse Gaussien et normal inverse Gaussien

Si $Y(t) \sim \text{InvGaussian}(t/\mu, t^2/\sigma^2)$, alors $Y = \{Y(t), t \geq 0\}$ est un processus inverse Gaussien.

Ce $Y(t)$ correspond au temps d'atteinte du niveau t par un MB avec dérive μ et variance σ^2 . $\mathbb{E}[Y(t)] = t/\mu$ et $\text{Var}[Y(t)] = t\sigma^2/\mu^3$.

Processus inverse Gaussien et normal inverse Gaussien

Si $Y(t) \sim \text{InvGaussian}(t/\mu, t^2/\sigma^2)$, alors $Y = \{Y(t), t \geq 0\}$ est un processus inverse Gaussien.

Ce $Y(t)$ correspond au temps d'atteinte du niveau t par un MB avec dérive μ et variance σ^2 . $\mathbb{E}[Y(t)] = t/\mu$ et $\text{Var}[Y(t)] = t\sigma^2/\mu^3$.

Ce processus a une trajectoire non décroissante. On peut donc l'utiliser comme subordonateur dans un autre processus de Lévy.

Si on fait cela pour un MB, cela donne processus normal inverse Gaussien (NIG) $\{Y(t), t \geq 0\}$, défini par

$$Y(t) = X(T(t)),$$

où T est un processus inverse Gaussien.

Séries Chronologiques

Pour un processus stochastique $\{Y_t, t \geq 0\}$, on définit:

$\mu_t = \mathbb{E}[Y_t]$, la fonction de moyenne,

$\sigma_t^2 = \text{Var}[Y_t]$, la fonction de variance,

$\text{Cov}[Y_t, Y_s]$, la fonction d'autocovariance,

$\rho_{t,s} = \text{Cov}[Y_t, Y_s] / \sigma_t \sigma_s$, la fonction d'autocorrelation.

Séries Chronologiques

Pour un processus stochastique $\{Y_t, t \geq 0\}$, on définit:

$\mu_t = \mathbb{E}[Y_t]$, la fonction de moyenne,

$\sigma_t^2 = \text{Var}[Y_t]$, la fonction de variance,

$\text{Cov}[Y_t, Y_s]$, la fonction d'autocovariance,

$\rho_{t,s} = \text{Cov}[Y_t, Y_s] / \sigma_t \sigma_s$, la fonction d'autocorrelation.

Le processus est faiblement stationnaire si $\mu_t \equiv \mu$, $\sigma_t^2 \equiv \sigma^2$ et $\rho_{t,t+s} \equiv \rho_s$ (autocorrélation de délai s) sont indépendants de t .

Séries Chronologiques

Pour un processus stochastique $\{Y_t, t \geq 0\}$, on définit:

$\mu_t = \mathbb{E}[Y_t]$, la fonction de moyenne,

$\sigma_t^2 = \text{Var}[Y_t]$, la fonction de variance,

$\text{Cov}[Y_t, Y_s]$, la fonction d'autocovariance,

$\rho_{t,s} = \text{Cov}[Y_t, Y_s] / \sigma_t \sigma_s$, la fonction d'autocorrelation.

Le processus est faiblement stationnaire si $\mu_t \equiv \mu$, $\sigma_t^2 \equiv \sigma^2$ et $\rho_{t,t+s} \equiv \rho_s$ (autocorrélation de délai s) sont indépendants de t .

Pour un tel processus, la loi de Y_t peut quand même dépendre de t .

Séries Chronologiques

Pour un processus stochastique $\{Y_t, t \geq 0\}$, on définit:

$\mu_t = \mathbb{E}[Y_t]$, la fonction de moyenne,

$\sigma_t^2 = \text{Var}[Y_t]$, la fonction de variance,

$\text{Cov}[Y_t, Y_s]$, la fonction d'autocovariance,

$\rho_{t,s} = \text{Cov}[Y_t, Y_s] / \sigma_t \sigma_s$, la fonction d'autocorrelation.

Le processus est **faiblement stationnaire** si $\mu_t \equiv \mu$, $\sigma_t^2 \equiv \sigma^2$ et $\rho_{t,t+s} \equiv \rho_s$ (autocorrélation de **délai** s) sont indépendants de t .

Pour un tel processus, la **loi** de Y_t peut quand même dépendre de t .

Le processus $\{Y_t, t \geq 0\}$ est **strictement stationnaire** si pour tout vecteur (t_1, \dots, t_k) fixé, la loi conjointe de $(Y_{t+t_1}, \dots, Y_{t+t_k})$ ne dépend pas de t .

Séries Chronologiques

Pour un processus stochastique $\{Y_t, t \geq 0\}$, on définit:

$\mu_t = \mathbb{E}[Y_t]$, la fonction de moyenne,

$\sigma_t^2 = \text{Var}[Y_t]$, la fonction de variance,

$\text{Cov}[Y_t, Y_s]$, la fonction d'autocovariance,

$\rho_{t,s} = \text{Cov}[Y_t, Y_s] / \sigma_t \sigma_s$, la fonction d'autocorrelation.

Le processus est **faiblement stationnaire** si $\mu_t \equiv \mu$, $\sigma_t^2 \equiv \sigma^2$ et $\rho_{t,t+s} \equiv \rho_s$ (autocorrélation de **délai** s) sont indépendants de t .

Pour un tel processus, la **loi** de Y_t peut quand même dépendre de t .

Le processus $\{Y_t, t \geq 0\}$ est **strictement stationnaire** si pour tout vecteur (t_1, \dots, t_k) fixé, la loi conjointe de $(Y_{t+t_1}, \dots, Y_{t+t_k})$ ne dépend pas de t .

En pratique, il est très difficile de modéliser cette loi conjointe. Souvent, on se restreint à une classe étroite de modèles pour lesquels il ne reste à spécifier que quelques paramètres (e.g., moyenne, variance et autocorrelations).

Une **série chronologique** est un processus stochastique en temps discret $\{Y_n, n \geq 0\}$.

Une **série chronologique** est un processus stochastique en temps discret $\{Y_n, n \geq 0\}$.

Exemples.

Modèle **AR**(p) (**autorégressif** d'ordre p):

$$Y_n = \phi_1 Y_{n-1} + \cdots + \phi_p Y_{n-p} + \theta_0 + \epsilon_n.$$

où les ϵ_n sont i.i.d. $N(0, \sigma^2)$.

Une **série chronologique** est un processus stochastique en temps discret $\{Y_n, n \geq 0\}$.

Exemples.

Modèle **AR**(p) (**autorégressif** d'ordre p):

$$Y_n = \phi_1 Y_{n-1} + \cdots + \phi_p Y_{n-p} + \theta_0 + \epsilon_n.$$

où les ϵ_n sont i.i.d. $N(0, \sigma^2)$.

Modèle **ARIMA**(p, d, q):

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d Y_n = \theta_0 + (1 - \theta_1 B - \cdots - \theta_q B^q) \epsilon_n,$$

où $B^r Y_n = Y_{n-r}$. Ici, tout suit la loi normale.

Modèles ARTA et VARTA

Même esprit que pour NORTA.

Pour un processus $\text{ARTA}(p)$ (“autoregressive to anything”) on spécifie la fonction de répartition marginale de Y_n , disons F , et les autocorrélations de délai $\leq p$.

Modèles ARTA et VARTA

Même esprit que pour NORTA.

Pour un processus $\text{ARTA}(p)$ (“autoregressive to anything”) on spécifie la fonction de répartition marginale de Y_n , disons F , et les autocorrélations de délai $\leq p$.

On pose $Y_n = F^{-1}[\Phi(Z_n)]$ où $\{Z_n, n \geq 0\}$ est $\text{AR}(p)$ avec les paramètres choisis pour atteindre les corrélations visées pour les Y_n .

Modèles ARTA et VARTA

Même esprit que pour NORTA.

Pour un processus $\text{ARTA}(p)$ (“autoregressive to anything”) on spécifie la fonction de répartition marginale de Y_n , disons F , et les autocorrélations de délai $\leq p$.

On pose $Y_n = F^{-1}[\Phi(Z_n)]$ où $\{Z_n, n \geq 0\}$ est $\text{AR}(p)$ avec les paramètres choisis pour atteindre les corrélations visées pour les Y_n .

Le processus $\text{VARTA}(p)$ (“vector autoregressive to anything”) est une version multivariée.

Modèles ARTA et VARTA

Même esprit que pour NORTA.

Pour un processus $\text{ARTA}(p)$ (“autoregressive to anything”) on spécifie la fonction de répartition marginale de Y_n , disons F , et les autocorrélations de délai $\leq p$.

On pose $Y_n = F^{-1}[\Phi(Z_n)]$ où $\{Z_n, n \geq 0\}$ est $\text{AR}(p)$ avec les paramètres choisis pour atteindre les corrélations visées pour les Y_n .

Le processus $\text{VARTA}(p)$ (“vector autoregressive to anything”) est une version multivariée.

Nelson et al. ont développé des méthodes d'estimation et des logiciels pour le cas où F est de la famille de Johnson.

Minification, maxification, TES

Idée: transformer des uniformes i.i.d. en uniformes corrélées,
et leur appliquer ensuite F^{-1} pour le F choisi.

Minification, maxification, TES

Idée: transformer des uniformes i.i.d. en uniformes corrélées, et leur appliquer ensuite F^{-1} pour le F choisi.

Transform-expand-sample (TES) (Livny et al. 1993).

Choisir des paramètres $-0.5 < L < R \leq 0.5$. $\{Z_n\}$ i.i.d. $U(0, 1)$.

Minification, maxification, TES

Idée: transformer des uniformes i.i.d. en uniformes corrélées, et leur appliquer ensuite F^{-1} pour le F choisi.

Transform-expand-sample (TES) (Livny et al. 1993).

Choisir des paramètres $-0.5 < L < R \leq 0.5$. $\{Z_n\}$ i.i.d. $U(0, 1)$.

$$U_0 = Z_0; \quad U_n = (U_{n-1} + L + (R - L)Z_n) \bmod 1.$$

$$Y_n = F^{-1}(U_n).$$

Minification, maxification, TES

Idée: transformer des uniformes i.i.d. en uniformes corrélées, et leur appliquer ensuite F^{-1} pour le F choisi.

Transform-expand-sample (TES) (Livny et al. 1993).

Choisir des paramètres $-0.5 < L < R \leq 0.5$. $\{Z_n\}$ i.i.d. $U(0, 1)$.

$U_0 = Z_0$; $U_n = (U_{n-1} + L + (R - L)Z_n) \bmod 1$.

$Y_n = F^{-1}(U_n)$. Rapide.

On dispose d'une formule pour calculer les ρ_j ; ils dépendent de R, L et F .

Minification, maxification, TES

Idée: transformer des uniformes i.i.d. en uniformes corrélées, et leur appliquer ensuite F^{-1} pour le F choisi.

Transform-expand-sample (TES) (Livny et al. 1993).

Choisir des paramètres $-0.5 < L < R \leq 0.5$. $\{Z_n\}$ i.i.d. $U(0, 1)$.

$U_0 = Z_0$; $U_n = (U_{n-1} + L + (R - L)Z_n) \bmod 1$.

$Y_n = F^{-1}(U_n)$. Rapide.

On dispose d'une formule pour calculer les ρ_j ; ils dépendent de R, L et F .

Une variante qui produit des autocorrélations dont le signe alterne.

Minification, maxification, TES

Idée: transformer des uniformes i.i.d. en uniformes corrélées, et leur appliquer ensuite F^{-1} pour le F choisi.

Transform-expand-sample (TES) (Livny et al. 1993).

Choisir des paramètres $-0.5 < L < R \leq 0.5$. $\{Z_n\}$ i.i.d. $U(0, 1)$.

$U_0 = Z_0$; $U_n = (U_{n-1} + L + (R - L)Z_n) \bmod 1$.

$Y_n = F^{-1}(U_n)$. Rapide.

On dispose d'une formule pour calculer les ρ_j ; ils dépendent de R, L et F .

Une variante qui produit des autocorrélations dont le signe alterne.

Les $|\rho_j|$ décroissent à un taux plus lent qu'une exponentielle.

Minification, maxification, TES

Idée: transformer des uniformes i.i.d. en uniformes corrélées, et leur appliquer ensuite F^{-1} pour le F choisi.

Transform-expand-sample (TES) (Livny et al. 1993).

Choisir des paramètres $-0.5 < L < R \leq 0.5$. $\{Z_n\}$ i.i.d. $U(0, 1)$.

$U_0 = Z_0$; $U_n = (U_{n-1} + L + (R - L)Z_n) \bmod 1$.

$Y_n = F^{-1}(U_n)$. Rapide.

On dispose d'une formule pour calculer les ρ_j ; ils dépendent de R, L et F .

Une variante qui produit des autocorrélations dont le signe alterne.

Les $|\rho_j|$ décroissent à un taux plus lent qu'une exponentielle.

Méthode de minification (Lewis and McKenzie 1991):

Paramètre c .

$U_n := c \cdot \min\{U_{n-1}, Z_{n-1}/(Z_{n-1} + c - 1)\}$.

Les ρ_j pour les U_n satisfont $\rho_j = c^{-j}$.

Minification, maxification, TES

Idée: transformer des uniformes i.i.d. en uniformes corrélées, et leur appliquer ensuite F^{-1} pour le F choisi.

Transform-expand-sample (TES) (Livny et al. 1993).

Choisir des paramètres $-0.5 < L < R \leq 0.5$. $\{Z_n\}$ i.i.d. $U(0, 1)$.

$U_0 = Z_0$; $U_n = (U_{n-1} + L + (R - L)Z_n) \bmod 1$.

$Y_n = F^{-1}(U_n)$. Rapide.

On dispose d'une formule pour calculer les ρ_j ; ils dépendent de R, L et F .

Une variante qui produit des autocorrélations dont le signe alterne.

Les $|\rho_j|$ décroissent à un taux plus lent qu'une exponentielle.

Méthode de minification (Lewis and McKenzie 1991):

Paramètre c .

$U_n := c \cdot \min\{U_{n-1}, Z_{n-1}/(Z_{n-1} + c - 1)\}$.

Les ρ_j pour les U_n satisfont $\rho_j = c^{-j}$.

Exemple: Effet sur l'attente dans une file; voir notes de cours.

Estimation de paramètres

On a n observations x_1, \dots, x_n venant d'une loi de densité $f_{\theta}(x)$ et on veut estimer θ (inconnu). Ici, x et θ peuvent être des vecteurs.

Estimation de paramètres

On a n observations x_1, \dots, x_n venant d'une loi de densité $f_{\theta}(x)$ et on veut estimer θ (inconnu). Ici, x et θ peuvent être des vecteurs.

Fonction de vraisemblance de l'échantillon:

$$L(\theta) = f_{\theta}(x_1) \cdots f_{\theta}(x_n).$$

Estimateur de vraisemblance maximale (EVM): c'est la valeur de θ qui maximise $L(\theta)$. Plusieurs propriétés intéressantes font que c'est la méthode de choix.

Estimation de paramètres

On a n observations x_1, \dots, x_n venant d'une loi de densité $f_{\theta}(x)$ et on veut estimer θ (inconnu). Ici, x et θ peuvent être des vecteurs.

Fonction de vraisemblance de l'échantillon:

$$L(\theta) = f_{\theta}(x_1) \cdots f_{\theta}(x_n).$$

Estimateur de vraisemblance maximale (EVM): c'est la valeur de θ qui maximise $L(\theta)$. Plusieurs propriétés intéressantes font que c'est la méthode de choix.

Par exemple, $\sqrt{n}(\hat{\theta}_n - \theta) \Rightarrow N(\mathbf{0}, n(\mathbf{I}(\theta))^{-1})$.

Estimation de paramètres

On a n observations x_1, \dots, x_n venant d'une loi de densité $f_{\theta}(x)$ et on veut estimer θ (inconnu). Ici, x et θ peuvent être des vecteurs.

Fonction de vraisemblance de l'échantillon:

$$L(\theta) = f_{\theta}(x_1) \cdots f_{\theta}(x_n).$$

Estimateur de vraisemblance maximale (EVM): c'est la valeur de θ qui maximise $L(\theta)$. Plusieurs propriétés intéressantes font que c'est la méthode de choix.

Par exemple, $\sqrt{n}(\hat{\theta}_n - \theta) \Rightarrow N(\mathbf{0}, n(\mathbf{I}(\theta))^{-1})$.

Autres méthodes: ajustement des moments (“moment matching”), moindres carrés,...

Exemple. Loi **Weibull** avec $\delta = 0$ et $\boldsymbol{\theta} = (\alpha, \lambda)$. On a

$$\begin{aligned} f(x) &= \alpha \lambda^\alpha x^{\alpha-1} e^{-\lambda x^\alpha} \quad \text{pour } x > 0, \\ L(\alpha, \lambda) &= \alpha^n \lambda^{n\alpha} (x_1 \cdots x_n)^{\alpha-1} e^{-\lambda^\alpha (x_1^\alpha + \cdots + x_n^\alpha)}, \\ \ln L(\alpha, \lambda) &= n \ln \alpha + n\alpha \ln \lambda + (\alpha - 1) \sum_{i=1}^n \ln x_i - \lambda^\alpha \sum_{i=1}^n x_i^\alpha. \end{aligned}$$

On regarde où la dérivée de $\ln L$ vaut zéro, car c'est plus facile que pour L .

Exemple. Loi **Weibull** avec $\delta = 0$ et $\boldsymbol{\theta} = (\alpha, \lambda)$. On a

$$\begin{aligned} f(x) &= \alpha \lambda^\alpha x^{\alpha-1} e^{-\lambda x^\alpha} \quad \text{pour } x > 0, \\ L(\alpha, \lambda) &= \alpha^n \lambda^{n\alpha} (x_1 \cdots x_n)^{\alpha-1} e^{-\lambda^\alpha (x_1^\alpha + \cdots + x_n^\alpha)}, \\ \ln L(\alpha, \lambda) &= n \ln \alpha + n\alpha \ln \lambda + (\alpha - 1) \sum_{i=1}^n \ln x_i - \lambda^\alpha \sum_{i=1}^n x_i^\alpha. \end{aligned}$$

On regarde où la dérivée de $\ln L$ vaut zéro, car c'est plus facile que pour L .
En mettant les dérivées partielles à zéro, on obtient

$$\begin{aligned} \frac{\partial \ln L(\alpha, \lambda)}{\partial \alpha} &= \frac{n}{\alpha} + n \ln \lambda + \sum_{i=1}^n \ln x_i - \sum_{i=1}^n (\lambda x_i)^\alpha \ln(\lambda x_i) = 0, \\ \frac{\partial \ln L(\alpha, \lambda)}{\partial \lambda} &= \frac{\alpha n}{\lambda} - \alpha \lambda^{\alpha-1} \sum_{i=1}^n x_i^\alpha = 0. \end{aligned}$$

La seconde équation permet d'écrire λ en fonction de α , puis on remplace dans la première et on la résoud numériquement.

Exemple. Loi **Weibull** avec $\delta = 0$ et $\boldsymbol{\theta} = (\alpha, \lambda)$. On a

$$\begin{aligned} f(x) &= \alpha \lambda^\alpha x^{\alpha-1} e^{-\lambda x^\alpha} \quad \text{pour } x > 0, \\ L(\alpha, \lambda) &= \alpha^n \lambda^{n\alpha} (x_1 \cdots x_n)^{\alpha-1} e^{-\lambda^\alpha (x_1^\alpha + \cdots + x_n^\alpha)}, \\ \ln L(\alpha, \lambda) &= n \ln \alpha + n\alpha \ln \lambda + (\alpha - 1) \sum_{i=1}^n \ln x_i - \lambda^\alpha \sum_{i=1}^n x_i^\alpha. \end{aligned}$$

On regarde où la dérivée de $\ln L$ vaut zéro, car c'est plus facile que pour L .
En mettant les dérivées partielles à zéro, on obtient

$$\begin{aligned} \frac{\partial \ln L(\alpha, \lambda)}{\partial \alpha} &= \frac{n}{\alpha} + n \ln \lambda + \sum_{i=1}^n \ln x_i - \sum_{i=1}^n (\lambda x_i)^\alpha \ln(\lambda x_i) = 0, \\ \frac{\partial \ln L(\alpha, \lambda)}{\partial \lambda} &= \frac{\alpha n}{\lambda} - \alpha \lambda^{\alpha-1} \sum_{i=1}^n x_i^\alpha = 0. \end{aligned}$$

La seconde équation permet d'écrire λ en fonction de α , puis on remplace dans la première et on la résoud numériquement.

Pour $\alpha = 1$, on obtient: $\hat{\lambda}_n = 1/\bar{x}_n = n/\sum_{i=1}^n x_i$.

Tests d'Ajustement

Voir Law et Kelton (2000) pour plus de détails.

On veut tester si la f.r. F s'ajuste bien aux données.

Tests d'Ajustement

Voir Law et Kelton (2000) pour plus de détails.

On veut tester si la f.r. F s'ajuste bien aux données.

Tests visuels, histogrammes.

Graphique P-P: points $((i - 0.5)/n, F(x_{(i)})) = (\hat{F}_n(x_{(i)}) - 0.5/n, F(x_{(i)}))$;

Tests d'Ajustement

Voir Law et Kelton (2000) pour plus de détails.

On veut tester si la f.r. F s'ajuste bien aux données.

Tests visuels, histogrammes.

Graphique P-P: points $((i - 0.5)/n, F(x_{(i)})) = (\hat{F}_n(x_{(i)}) - 0.5/n, F(x_{(i)}))$;

Graphique Q-Q: points $(F^{-1}((i - 0.5)/n), x_{(i)})$.

Tests d'Ajustement

Voir Law et Kelton (2000) pour plus de détails.

On veut tester si la f.r. F s'ajuste bien aux données.

Tests visuels, histogrammes.

Graphique P-P: points $((i - 0.5)/n, F(x_{(i)})) = (\hat{F}_n(x_{(i)}) - 0.5/n, F(x_{(i)}))$;

Graphique Q-Q: points $(F^{-1}((i - 0.5)/n), x_{(i)})$.

Le P-P plot détecte les différences davantage au centre, le Q-Q plot davantage dans les queues.

Tests d'Ajustement

Voir Law et Kelton (2000) pour plus de détails.

On veut tester si la f.r. F s'ajuste bien aux données.

Tests visuels, histogrammes.

Graphique P-P: points $((i - 0.5)/n, F(x_{(i)})) = (\hat{F}_n(x_{(i)}) - 0.5/n, F(x_{(i)}))$;

Graphique Q-Q: points $(F^{-1}((i - 0.5)/n), x_{(i)})$.

Le P-P plot détecte les différences davantage au centre, le Q-Q plot davantage dans les queues.

Chi-deux, Kolmogorov-Smirnov, Anderson Darling.

Tests d'Ajustement

Voir Law et Kelton (2000) pour plus de détails.

On veut tester si la f.r. F s'ajuste bien aux données.

Tests visuels, histogrammes.

Graphique P-P: points $((i - 0.5)/n, F(x_{(i)})) = (\hat{F}_n(x_{(i)}) - 0.5/n, F(x_{(i)}))$;

Graphique Q-Q: points $(F^{-1}((i - 0.5)/n), x_{(i)})$.

Le P-P plot détecte les différences davantage au centre, le Q-Q plot davantage dans les queues.

Chi-deux, Kolmogorov-Smirnov, Anderson Darling.

Kolmogorov-Smirnov:

$$\begin{aligned} D_n &= \sup_{-\infty < x < \infty} |\hat{F}_n(x) - F(x)| \\ &= \max \left(\max_{1 \leq i \leq n} [i/n - u_{(i)}], \max_{1 \leq i \leq n} [u_{(i)} - (i - 1)/n] \right). \end{aligned}$$

On rejette \mathcal{H}_0 : "la bonne loi est F " si D_n est trop grand.