

Moyenne sur horizon infini, état stationnaire

Soit $\{C_i, i \geq 1\}$ un processus faiblement stationnaire;

$\mathbb{E}[C_i] = \mu$, $\text{Var}[C_i] = \sigma^2 > 0$, autocorrelations $\rho_k = \rho_{i,i+k} = \text{Corr}[C_i, C_{i+k}]$.

Moyenne sur horizon infini, état stationnaire

Soit $\{C_i, i \geq 1\}$ un processus faiblement stationnaire;

$\mathbb{E}[C_i] = \mu$, $\text{Var}[C_i] = \sigma^2 > 0$, autocorrelations $\rho_k = \rho_{i,i+k} = \text{Corr}[C_i, C_{i+k}]$.

On veut estimer μ par la moyenne empirique

$$\bar{C}_n = \frac{1}{n} \sum_{i=1}^n C_i.$$

Moyenne sur horizon infini, état stationnaire

Soit $\{C_i, i \geq 1\}$ un processus faiblement stationnaire;

$\mathbb{E}[C_i] = \mu$, $\text{Var}[C_i] = \sigma^2 > 0$, autocorrelations $\rho_k = \rho_{i,i+k} = \text{Corr}[C_i, C_{i+k}]$.

On veut estimer μ par la moyenne empirique

$$\bar{C}_n = \frac{1}{n} \sum_{i=1}^n C_i.$$

On a $\mathbb{E}[\bar{C}_n] = \mu$ et

$$\text{Var}[\bar{C}_n] = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \text{Cov}[C_i, C_j] = \frac{\sigma^2}{n} + \frac{2\sigma^2}{n^2} \sum_{i=1}^n \sum_{j=i+1}^n \rho_{i,j} = \frac{\sigma^2(1 + \gamma_n)}{n}$$

où

$$\gamma_n = \frac{2}{n} \sum_{i=1}^n \sum_{j=i+1}^n \rho_{i,j} = \frac{2}{n} \sum_{i=1}^n \sum_{j=i+1}^n \rho_{j-i} = \frac{2}{n} \sum_{k=1}^{n-1} (n-k) \rho_k.$$

Soit

$$\gamma = \lim_{n \rightarrow \infty} \gamma_n = 2 \sum_{k=1}^{\infty} \rho_k$$

si la limite existe. Si $\gamma < \infty$, alors

$$\text{Var}[\bar{C}_n] = \frac{\sigma^2(1 + \gamma_n)}{n} = \frac{\sigma_\infty^2}{n} + o(1/n)$$

où

$$\sigma_\infty^2 = \lim_{n \rightarrow \infty} \sigma_n^2 = \lim_{n \rightarrow \infty} n \text{Var}[\bar{C}_n] = \sigma^2(1 + \gamma).$$

Soit

$$\gamma = \lim_{n \rightarrow \infty} \gamma_n = 2 \sum_{k=1}^{\infty} \rho_k$$

si la limite existe. Si $\gamma < \infty$, alors

$$\text{Var}[\bar{C}_n] = \frac{\sigma^2(1 + \gamma_n)}{n} = \frac{\sigma_\infty^2}{n} + o(1/n)$$

où

$$\sigma_\infty^2 = \lim_{n \rightarrow \infty} \sigma_n^2 = \lim_{n \rightarrow \infty} n \text{Var}[\bar{C}_n] = \sigma^2(1 + \gamma).$$

La variance est gonflée par un facteur $\sigma_\infty^2/\sigma^2 = 1 + \gamma$ par rapport au cas des C_i indépendants. On peut voir γ comme une mesure globale d'autocorrélation.

Soit

$$\gamma = \lim_{n \rightarrow \infty} \gamma_n = 2 \sum_{k=1}^{\infty} \rho_k$$

si la limite existe. Si $\gamma < \infty$, alors

$$\text{Var}[\bar{C}_n] = \frac{\sigma^2(1 + \gamma_n)}{n} = \frac{\sigma_\infty^2}{n} + o(1/n)$$

où

$$\sigma_\infty^2 = \lim_{n \rightarrow \infty} \sigma_n^2 = \lim_{n \rightarrow \infty} n \text{Var}[\bar{C}_n] = \sigma^2(1 + \gamma).$$

La **variance est gonflée par un facteur** $\sigma_\infty^2/\sigma^2 = 1 + \gamma$ par rapport au cas des C_i indépendants. On peut voir γ comme une **mesure globale d'autocorrélation**.

Si \bar{C}_n obéit à un TLC, il doit avoir la forme

$$\frac{\sqrt{n}(\bar{C}_n - \mu)}{\sigma_\infty} \Rightarrow N(0, 1) \text{ quand } n \rightarrow \infty.$$

Si $\hat{\sigma}_n^2$ est un estimateur consistant pour σ_∞^2 , on s'attend à avoir le TLC:

$$\frac{\sqrt{n}(\bar{C}_n - \mu)}{\hat{\sigma}_n} \Rightarrow \frac{\sqrt{n}(\bar{C}_n - \mu)}{\sigma_\infty} \Rightarrow N(0, 1) \text{ quand } n \rightarrow \infty.$$

Si $\hat{\sigma}_n^2$ est un estimateur consistant pour σ_∞^2 , on s'attend à avoir le TLC:

$$\frac{\sqrt{n}(\bar{C}_n - \mu)}{\hat{\sigma}_n} \Rightarrow \frac{\sqrt{n}(\bar{C}_n - \mu)}{\sigma_\infty} \Rightarrow N(0, 1) \text{ quand } n \rightarrow \infty.$$

- (a) Sous quelles conditions a-t-on ce TLC?
- (b) Comment construire $\hat{\sigma}_n^2$?

Si $\hat{\sigma}_n^2$ est un estimateur consistant pour σ_∞^2 , on s'attend à avoir le TLC:

$$\frac{\sqrt{n}(\bar{C}_n - \mu)}{\hat{\sigma}_n} \Rightarrow \frac{\sqrt{n}(\bar{C}_n - \mu)}{\sigma_\infty} \Rightarrow N(0, 1) \text{ quand } n \rightarrow \infty.$$

- (a) Sous quelles conditions a-t-on ce TLC?
- (b) Comment construire $\hat{\sigma}_n^2$?

Pour répondre à (a), relaxons maintenant l'hypothèse de stationnarité.
 Nous allons travailler dans le contexte où tient un **TLC fonctionnel** (TLCF), qui implique une stationnarité asymptotique.

Si $\hat{\sigma}_n^2$ est un estimateur consistant pour σ_∞^2 , on s'attend à avoir le TLC:

$$\frac{\sqrt{n}(\bar{C}_n - \mu)}{\hat{\sigma}_n} \Rightarrow \frac{\sqrt{n}(\bar{C}_n - \mu)}{\sigma_\infty} \Rightarrow N(0, 1) \text{ quand } n \rightarrow \infty.$$

- (a) Sous quelles conditions a-t-on ce TLC?
- (b) Comment construire $\hat{\sigma}_n^2$?

Pour répondre à (a), relaxons maintenant l'hypothèse de stationnarité. Nous allons travailler dans le contexte où tient un **TLC fonctionnel** (TLCF), qui implique une stationnarité asymptotique.

Pour des constantes μ et σ_∞ , et pour chaque n , on définit le processus

$$W_n(t) = \frac{[nt](\bar{C}_{[nt]} - \mu)}{\sigma_\infty \sqrt{n}}, \quad 0 \leq t \leq 1,$$

où $\bar{C}_0 = \mu$.

Propriété TLCF: $W_n \Rightarrow W$ où W est un mouvement Brownien standard.

De cette propriété découle bien sûr le **TLC**:

$$\frac{\sqrt{n}(\bar{C}_n - \mu)}{\hat{\sigma}_n} \Rightarrow \frac{\sqrt{n}(\bar{C}_n - \mu)}{\sigma_\infty} = W_n(1) \Rightarrow W(1) \sim N(0, 1),$$

que l'on peut utiliser pour construire un IC pour μ .

De cette propriété découle bien sûr le **TLC**:

$$\frac{\sqrt{n}(\bar{C}_n - \mu)}{\hat{\sigma}_n} \Rightarrow \frac{\sqrt{n}(\bar{C}_n - \mu)}{\sigma_\infty} = W_n(1) \Rightarrow W(1) \sim N(0, 1),$$

que l'on peut utiliser pour construire un IC pour μ .

Mais quand tient-elle et comment la vérifier?

En gros, il suffit que la dépendance entre $\{C_1, \dots, C_j\}$ et $\{C_{j+n}, C_{j+n+1}, \dots\}$ s'estompe assez rapidement lorsque $n \rightarrow \infty$.

En gros, il suffit que la dépendance entre $\{C_1, \dots, C_j\}$ et $\{C_{j+n}, C_{j+n+1}, \dots\}$ s'estompe assez rapidement lorsque $n \rightarrow \infty$.

Plus précisément (conditions classiques):

Définition. Pour $1 \leq i \leq j \leq \infty$, soit $\mathcal{F}_{i,j}$ l'information que l'on peut déduire en observant C_i, \dots, C_j .

En gros, il suffit que la dépendance entre $\{C_1, \dots, C_j\}$ et $\{C_{j+n}, C_{j+n+1}, \dots\}$ s'estompe assez rapidement lorsque $n \rightarrow \infty$.

Plus précisément (conditions classiques):

Définition. Pour $1 \leq i \leq j \leq \infty$, soit $\mathcal{F}_{i,j}$ l'information que l'on peut déduire en observant C_i, \dots, C_j .

$\{C_j, j \geq 1\}$ est “**phi-mixing**” s'il existe une suite de réels $\{\varphi_n, n \geq 1\}$ telle que $\varphi_n \rightarrow 0$, et pour tous $j \geq 1, n \geq 1$, $E_1 \in \mathcal{F}_{1,j}$ tel que $\mathbb{P}[E_1] > 0$, et $E_2 \in \mathcal{F}_{j+n,\infty}$, on a

$$|\mathbb{P}[E_2 \mid E_1] - \mathbb{P}[E_2]| \leq \varphi_n.$$

En gros, il suffit que la dépendance entre $\{C_1, \dots, C_j\}$ et $\{C_{j+n}, C_{j+n+1}, \dots\}$ s'estompe assez rapidement lorsque $n \rightarrow \infty$.

Plus précisément (conditions classiques):

Définition. Pour $1 \leq i \leq j \leq \infty$, soit $\mathcal{F}_{i,j}$ l'information que l'on peut déduire en observant C_i, \dots, C_j .

$\{C_j, j \geq 1\}$ est “**phi-mixing**” s'il existe une suite de réels $\{\varphi_n, n \geq 1\}$ telle que $\varphi_n \rightarrow 0$, et pour tous $j \geq 1, n \geq 1$, $E_1 \in \mathcal{F}_{1,j}$ tel que $\mathbb{P}[E_1] > 0$, et $E_2 \in \mathcal{F}_{j+n,\infty}$, on a

$$|\mathbb{P}[E_2 \mid E_1] - \mathbb{P}[E_2]| \leq \varphi_n.$$

Théorème. Si $\{C_j, j \geq 1\}$ est **fortement stationnaire et phi-mixing**, avec $\sum_{n=1}^{\infty} \sqrt{\varphi_n} < \infty$, alors $\sum_{k=1}^{\infty} |\rho_k| < \infty$, $\gamma < \infty$, et le TLCF tient.

En gros, il suffit que la dépendance entre $\{C_1, \dots, C_j\}$ et $\{C_{j+n}, C_{j+n+1}, \dots\}$ s'estompe assez rapidement lorsque $n \rightarrow \infty$.

Plus précisément (conditions classiques):

Définition. Pour $1 \leq i \leq j \leq \infty$, soit $\mathcal{F}_{i,j}$ l'information que l'on peut déduire en observant C_i, \dots, C_j .

$\{C_j, j \geq 1\}$ est “**phi-mixing**” s'il existe une suite de réels $\{\varphi_n, n \geq 1\}$ telle que $\varphi_n \rightarrow 0$, et pour tous $j \geq 1, n \geq 1$, $E_1 \in \mathcal{F}_{1,j}$ tel que $\mathbb{P}[E_1] > 0$, et $E_2 \in \mathcal{F}_{j+n,\infty}$, on a

$$|\mathbb{P}[E_2 \mid E_1] - \mathbb{P}[E_2]| \leq \varphi_n.$$

Théorème. Si $\{C_j, j \geq 1\}$ est **fortement stationnaire et phi-mixing**, avec $\sum_{n=1}^{\infty} \sqrt{\varphi_n} < \infty$, alors $\sum_{k=1}^{\infty} |\rho_k| < \infty$, $\gamma < \infty$, et le TLCF tient.

Cette condition n'est pas facile à vérifier directement.

Le TLCF tient aussi sous des conditions un peu moins fortes que cela.

Par exemple, la stationnarité peut n'être qu'asymptotique.

Conditions de Lyapunov. Conditions plus facilement vérifiables.

Conditions de Lyapunov. Conditions plus facilement vérifiables.

Plusieurs types: celle utilisée ici est plutôt forte.

Conditions de Lyapunov. Conditions plus facilement vérifiables.

Plusieurs types: celle utilisée ici est plutôt forte.

Contexte: chaîne de Markov (CM) en temps discret $\{Y_j, j \geq 0\}$ sur un espace \mathbb{S} .

Fonction de coût h : Dans l'état Y_j , on paye $C_j = h(Y_j)$, pour $j \geq 1$.

μ est le coût moyen par étape (e.g., par événement).

Conditions de Lyapunov. Conditions plus facilement vérifiables.

Plusieurs types: celle utilisée ici est plutôt forte.

Contexte: chaîne de Markov (CM) en temps discret $\{Y_j, j \geq 0\}$ sur un espace \mathbb{S} .

Fonction de coût h : Dans l'état Y_j , on paye $C_j = h(Y_j)$, pour $j \geq 1$.

μ est le coût moyen par étape (e.g., par événement).

Définition. La CM $\{Y_j, j \geq 0\}$ est v -uniformément ergodique, pour $v : \mathbb{S} \rightarrow [1, \infty)$, s'il y a une mesure de probabilité π sur \mathbb{S} telle que

$$\lim_{n \rightarrow \infty} \sup_{y \in \mathbb{S}} \sup_{h: |h| \leq v} (\mathbb{E}[h(Y_n) \mid Y_0 = y] - \mathbb{E}_\pi[h(Y)]) = 0.$$

La fonction v s'appelle **fonction test** ou **fonction de Lyapunov**.

Théorème. Les conditions (de Lyapunov) suivantes sont **suffisantes** pour qu'une CM non périodique $\{Y_j, j \geq 0\}$ soit v -uniformément ergodique:

Théorème. Les conditions (de Lyapunov) suivantes sont **suffisantes** pour qu'une CM non périodique $\{Y_j, j \geq 0\}$ soit v -uniformément ergodique:

Il existe $B \subseteq \mathbb{S}$, $0 < a < 1$, $c > 0$, $\delta > 0$, un entier $m \geq 1$, une loi de probabilité Q sur \mathbb{S} , et une fonction $v : \mathbb{S} \rightarrow [1, \infty)$ tels que

$$\begin{aligned} \mathbb{P}[Y_m \in \cdot \mid Y_0 = y] &\geq \delta Q(\cdot) \quad \text{pour tout } y \in B \text{ et} \\ \mathbb{E}[v(Y_1) \mid Y_0 = y] &\leq av(y) + c\mathbb{I}[y \in B] \quad \text{pour tout } y \in \mathbb{S}. \end{aligned}$$

Théorème. Les conditions (de Lyapunov) suivantes sont **suffisantes** pour qu'une CM non périodique $\{Y_j, j \geq 0\}$ soit v -uniformément ergodique:

Il existe $B \subseteq \mathbb{S}$, $0 < a < 1$, $c > 0$, $\delta > 0$, un entier $m \geq 1$, une loi de probabilité Q sur \mathbb{S} , et une fonction $v : \mathbb{S} \rightarrow [1, \infty)$ tels que

$$\begin{aligned} \mathbb{P}[Y_m \in \cdot \mid Y_0 = y] &\geq \delta Q(\cdot) \quad \text{pour tout } y \in B \text{ et} \\ \mathbb{E}[v(Y_1) \mid Y_0 = y] &\leq av(y) + c\mathbb{I}[y \in B] \quad \text{pour tout } y \in \mathbb{S}. \end{aligned}$$

On peut voir B comme le “centre” (région d'attraction) de l'espace d'états \mathbb{S} et v comme une fonction de pénalité qui augmente lorsqu'on s'écarte du centre. Les conditions garantissent que l'on est toujours attiré vers le centre. Ainsi, le système est stable.

Théorème. Les conditions (de Lyapunov) suivantes sont **suffisantes** pour qu'une CM non périodique $\{Y_j, j \geq 0\}$ soit v -uniformément ergodique:

Il existe $B \subseteq \mathbb{S}$, $0 < a < 1$, $c > 0$, $\delta > 0$, un entier $m \geq 1$, une loi de probabilité Q sur \mathbb{S} , et une fonction $v : \mathbb{S} \rightarrow [1, \infty)$ tels que

$$\begin{aligned} \mathbb{P}[Y_m \in \cdot \mid Y_0 = y] &\geq \delta Q(\cdot) \quad \text{pour tout } y \in B \text{ et} \\ \mathbb{E}[v(Y_1) \mid Y_0 = y] &\leq av(y) + c\mathbb{I}[y \in B] \quad \text{pour tout } y \in \mathbb{S}. \end{aligned}$$

On peut voir B comme le “centre” (région d'attraction) de l'espace d'états \mathbb{S} et v comme une fonction de pénalité qui augmente lorsqu'on s'écarte du centre. Les conditions garantissent que l'on est toujours attiré vers le centre. Ainsi, le système est stable.

Exemple: Pour $W_{j+1} = \max(0, W_j + S_j - A_j)$ (Lindley), si $\mathbb{E}[S_j - A_j] < 0$, on a $Y_j \equiv W_j$ et on peut prendre $v(y) = e^{\gamma y}$ pour un γ bien choisi.

Théorème. Si la CM $\{Y_j, j \geq 0\}$ est v -uniformément ergodique, alors

- (i) Elle possède une loi stationnaire unique, π , sur \mathbb{S} .

Théorème. Si la CM $\{Y_j, j \geq 0\}$ est v -uniformément ergodique, alors

- (i) Elle possède une loi stationnaire unique, π , sur \mathbb{S} .
- (ii) Loi forte des grands nombres. Si $h : \mathbb{S} \rightarrow \mathbb{R}$, $\mathbb{E}_\pi[|h(Y)|] < \infty$, $\mu = \mathbb{E}_\pi[h(Y)]$, et $C_j = h(Y_j)$, alors

$$\lim_{n \rightarrow \infty} \bar{C}_n \xrightarrow{\text{p.s.}} \mu.$$

Théorème. Si la CM $\{Y_j, j \geq 0\}$ est v -uniformément ergodique, alors

(i) Elle possède une **loi stationnaire** unique, π , sur \mathbb{S} .

(ii) **Loi forte des grands nombres.** Si $h : \mathbb{S} \rightarrow \mathbb{R}$, $\mathbb{E}_\pi[|h(Y)|] < \infty$, $\mu = \mathbb{E}_\pi[h(Y)]$, et $C_j = h(Y_j)$, alors

$$\lim_{n \rightarrow \infty} \bar{C}_n \xrightarrow{\text{p.s.}} \mu.$$

(iii) **TLCF.** Si $|h(y)| \leq v(y)$ pour tout $y \in \mathbb{S}$ et $\sigma_\infty > 0$, alors $\{C_j, j \geq 1\}$ obéit au TLCF.

Théorème. Si la CM $\{Y_j, j \geq 0\}$ est v -uniformément ergodique, alors

(i) Elle possède une **loi stationnaire** unique, π , sur \mathbb{S} .

(ii) **Loi forte des grands nombres.** Si $h : \mathbb{S} \rightarrow \mathbb{R}$, $\mathbb{E}_\pi[|h(Y)|] < \infty$, $\mu = \mathbb{E}_\pi[h(Y)]$, et $C_j = h(Y_j)$, alors

$$\lim_{n \rightarrow \infty} \bar{C}_n \xrightarrow{\text{p.s.}} \mu.$$

(iii) **TLCF.** Si $|h(y)| \leq v(y)$ pour tout $y \in \mathbb{S}$ et $\sigma_\infty > 0$, alors $\{C_j, j \geq 1\}$ obéit au TLCF.

(iv) Convergence du **biais** en $\mathcal{O}(1/n)$. Sous les conditions de (iii), si l'état initial Y_0 suit la loi π_0 sur \mathbb{S} et si $\mathbb{E}_{\pi_0}[v(Y_0)] < \infty$, alors

$$\mathbb{E}_{\pi_0} [|\bar{C}_n - \mu|] = \kappa/n + \mathcal{O}(\alpha^n)$$

pour un $\kappa < \infty$ et $\alpha < 1$.

Temps continu.

Pour un processus $\{C(t), t \geq 0\}$, on pose

$$\begin{aligned}\bar{C}(t) &= \frac{1}{t} \int_0^t C(s) ds, & t \geq 0, \\ W_n(t) &= \frac{nt(\bar{C}(nt) - \mu)}{\sigma_\infty \sqrt{n}}, & 0 \leq t \leq 1,\end{aligned}$$

et le TLCF ne change pas.

Temps continu.

Pour un processus $\{C(t), t \geq 0\}$, on pose

$$\begin{aligned}\bar{C}(t) &= \frac{1}{t} \int_0^t C(s) ds, & t \geq 0, \\ W_n(t) &= \frac{nt(\bar{C}(nt) - \mu)}{\sigma_\infty \sqrt{n}}, & 0 \leq t \leq 1,\end{aligned}$$

et le TLCF ne change pas.

Pour le “phi-mixing”, on remplace φ_n par $\varphi(t)$, $t \geq 0$, etc.

Temps continu.

Pour un processus $\{C(t), t \geq 0\}$, on pose

$$\begin{aligned}\bar{C}(t) &= \frac{1}{t} \int_0^t C(s) ds, & t \geq 0, \\ W_n(t) &= \frac{nt(\bar{C}(nt) - \mu)}{\sigma_\infty \sqrt{n}}, & 0 \leq t \leq 1,\end{aligned}$$

et le TLCF ne change pas.

Pour le “phi-mixing”, on remplace φ_n par $\varphi(t)$, $t \geq 0$, etc.

On remplace σ_n^2 par $\sigma^2(t) = t \text{Var}[\bar{C}(t)]$, et on a

$$\sigma_\infty^2 = \lim_{t \rightarrow \infty} \sigma^2(t) = 2 \int_0^\infty \text{Cov}[C(0), C(t)] dt.$$

Traitement semblable au cas discret.

Coût moyen par unité de temps

C_i = coût payé lors de l'événement i ;

$N(t)$ = nombre d'événements durant $[0, t]$.

Coût moyen par unité de temps

C_i = coût payé lors de l'événement i ;

$N(t)$ = nombre d'événements durant $[0, t]$.

Supposons que $N(t)/t \rightarrow \lambda$ lorsque $t \rightarrow \infty$.

Si $\bar{C}_n \rightarrow \mu$, alors le coût moyen par unité de temps est $\bar{v} = \lambda\mu$.

Coût moyen par unité de temps

C_i = coût payé lors de l'événement i ;

$N(t)$ = nombre d'événements durant $[0, t]$.

Supposons que $N(t)/t \rightarrow \lambda$ lorsque $t \rightarrow \infty$.

Si $\bar{C}_n \rightarrow \mu$, alors le coût moyen par unité de temps est $\bar{v} = \lambda\mu$.

Si on pose

$$\tilde{\sigma}_{\infty}^2 = \lambda\sigma_{\infty}^2,$$

on a le TLC:

Théorème. Si notre TLCHF tient, $N(t)/t \rightarrow \lambda$ où $0 < \lambda < \infty$, et $\hat{\sigma}_t^2 \rightarrow \tilde{\sigma}_{\infty}^2$, en probabilité, alors pour $t \rightarrow \infty$,

$$\begin{aligned} \frac{\sqrt{t}[V_{N(t)}/t - \bar{v}]}{\hat{\sigma}_t} &\Rightarrow \frac{\sqrt{N(t)/\lambda}[\lambda V_{N(t)}/N(t) - \lambda\mu]}{\tilde{\sigma}_{\infty}} \\ &= \frac{\sqrt{N(t)}[V_{N(t)}/N(t) - \mu]}{\sigma_{\infty}} \\ &\Rightarrow N(0, 1). \end{aligned}$$

Example: Processus AR(1).

$$C_{j+1} = \mu + a_1(C_j - \mu) + \epsilon_j, \quad j \geq 1,$$

où $-1 < a_1 < 1$, et les ϵ_j i.i.d. $N(0, \sigma_\epsilon^2)$.

Example: Processus AR(1).

$$C_{j+1} = \mu + a_1(C_j - \mu) + \epsilon_j, \quad j \geq 1,$$

où $-1 < a_1 < 1$, et les ϵ_j i.i.d. $N(0, \sigma_\epsilon^2)$.

Si $C_1 \sim N(\mu, \sigma^2)$ où $\sigma^2 = \sigma_\epsilon^2 / (1 - a_1)^2$, alors le processus est strictement stationnaire, $C_j \sim N(\mu, \sigma^2)$, $\rho_j = a_1^j$, et $\sigma_\infty^2 = \sigma^2(1 + a_1)/(1 - a_1)$.

Example: Processus AR(1).

$$C_{j+1} = \mu + a_1(C_j - \mu) + \epsilon_j, \quad j \geq 1,$$

où $-1 < a_1 < 1$, et les ϵ_j i.i.d. $N(0, \sigma_\epsilon^2)$.

Si $C_1 \sim N(\mu, \sigma^2)$ où $\sigma^2 = \sigma_\epsilon^2 / (1 - a_1)^2$, alors le processus est strictement stationnaire, $C_j \sim N(\mu, \sigma^2)$, $\rho_j = a_1^j$, et $\sigma_\infty^2 = \sigma^2(1 + a_1)/(1 - a_1)$.

On a $\sigma_\infty^2 / \sigma^2 \rightarrow \infty$ quand $a_1 \rightarrow 1$.

Example: Processus AR(1).

$$C_{j+1} = \mu + a_1(C_j - \mu) + \epsilon_j, \quad j \geq 1,$$

où $-1 < a_1 < 1$, et les ϵ_j i.i.d. $N(0, \sigma_\epsilon^2)$.

Si $C_1 \sim N(\mu, \sigma^2)$ où $\sigma^2 = \sigma_\epsilon^2 / (1 - a_1)^2$, alors le processus est strictement stationnaire, $C_j \sim N(\mu, \sigma^2)$, $\rho_j = a_1^j$, et $\sigma_\infty^2 = \sigma^2(1 + a_1)/(1 - a_1)$.

On a $\sigma_\infty^2 / \sigma^2 \rightarrow \infty$ quand $a_1 \rightarrow 1$.

Si $a_1 < 0$, le signe des corrélations alterne et $\sigma_\infty^2 < \sigma^2$.

Example: Processus AR(1).

$$C_{j+1} = \mu + a_1(C_j - \mu) + \epsilon_j, \quad j \geq 1,$$

où $-1 < a_1 < 1$, et les ϵ_j i.i.d. $N(0, \sigma_\epsilon^2)$.

Si $C_1 \sim N(\mu, \sigma^2)$ où $\sigma^2 = \sigma_\epsilon^2 / (1 - a_1)^2$, alors le processus est strictement stationnaire, $C_j \sim N(\mu, \sigma^2)$, $\rho_j = a_1^j$, et $\sigma_\infty^2 = \sigma^2(1 + a_1)/(1 - a_1)$.

On a $\sigma_\infty^2 / \sigma^2 \rightarrow \infty$ quand $a_1 \rightarrow 1$.

Si $a_1 < 0$, le signe des corrélations alterne et $\sigma_\infty^2 < \sigma^2$.

Exercice: Vérifiez les conditions de Lyapunov.

Exemple: File $M/M/1$. Taux d'arrivée λ , taux de service μ .

12

$$W_{i+1} = \max(0, W_i + S_i - A_i).$$

Exemple: File $M/M/1$. Taux d'arrivée λ , taux de service μ .

$$W_{i+1} = \max(0, W_i + S_i - A_i).$$

La loi d'équilibre des temps d'attentes $\{W_j, j \geq 1\}$ est donnée par $F_w(x) = \mathbb{P}[W_j \leq x] = 1 - \rho e^{-(\mu-\lambda)x}$ pour $x \geq 0$, où $\rho = \lambda/\mu$.

Exemple: File $M/M/1$. Taux d'arrivée λ , taux de service μ .

$$W_{i+1} = \max(0, W_i + S_i - A_i).$$

La loi d'équilibre des temps d'attentes $\{W_j, j \geq 1\}$ est donnée par $F_w(x) = \mathbb{P}[W_j \leq x] = 1 - \rho e^{-(\mu-\lambda)x}$ pour $x \geq 0$, où $\rho = \lambda/\mu$.

Si W_1 suit cette loi, chaque W_j aussi et on a

$$w = \mathbb{E}[W_j] = \frac{\rho}{\mu(1-\rho)} \quad \text{et} \quad \sigma^2 = \text{Var}[W_j] = \frac{\rho(2-\rho)}{\mu^2(1-\rho)^2}.$$

(Important: ce n'est pas équivalent à démarrer le processus à $W_1 = w$.)

Exemple: File $M/M/1$. Taux d'arrivée λ , taux de service μ .

$$W_{i+1} = \max(0, W_i + S_i - A_i).$$

La loi d'équilibre des temps d'attentes $\{W_j, j \geq 1\}$ est donnée par $F_w(x) = \mathbb{P}[W_j \leq x] = 1 - \rho e^{-(\mu-\lambda)x}$ pour $x \geq 0$, où $\rho = \lambda/\mu$.

Si W_1 suit cette loi, chaque W_j aussi et on a

$$w = \mathbb{E}[W_j] = \frac{\rho}{\mu(1-\rho)} \quad \text{et} \quad \sigma^2 = \text{Var}[W_j] = \frac{\rho(2-\rho)}{\mu^2(1-\rho)^2}.$$

(Important: ce n'est pas équivalent à démarrer le processus à $W_1 = w$.)

On a aussi une formule pour ρ_j , $\lim_{j \rightarrow \infty} \rho_{j+1}/\rho_j = 4\rho/(1+\rho)^2 < 1$, et

$$\sigma_\infty^2 = \frac{\rho(\rho^3 - 4\rho^2 + 5\rho + 2)}{\mu^2(1-\rho)^4}.$$

Exemple: File $M/M/1$. Taux d'arrivée λ , taux de service μ .

$$W_{i+1} = \max(0, W_i + S_i - A_i).$$

La loi d'équilibre des temps d'attentes $\{W_j, j \geq 1\}$ est donnée par $F_w(x) = \mathbb{P}[W_j \leq x] = 1 - \rho e^{-(\mu-\lambda)x}$ pour $x \geq 0$, où $\rho = \lambda/\mu$.

Si W_1 suit cette loi, chaque W_j aussi et on a

$$w = \mathbb{E}[W_j] = \frac{\rho}{\mu(1-\rho)} \quad \text{et} \quad \sigma^2 = \text{Var}[W_j] = \frac{\rho(2-\rho)}{\mu^2(1-\rho)^2}.$$

(Important: ce n'est pas équivalent à démarrer le processus à $W_1 = w$.)

On a aussi une formule pour ρ_j , $\lim_{j \rightarrow \infty} \rho_{j+1}/\rho_j = 4\rho/(1+\rho)^2 < 1$, et

$$\sigma_\infty^2 = \frac{\rho(\rho^3 - 4\rho^2 + 5\rho + 2)}{\mu^2(1-\rho)^4}.$$

Pour ρ proche de 1, $\sigma_\infty^2 \approx 4\rho/[\mu^2(1-\rho)^4] \approx 4\sigma^2/(1-\rho)^2 \gg \sigma^2$.

Exemple: File $M/M/1$. Taux d'arrivée λ , taux de service μ .

$$W_{i+1} = \max(0, W_i + S_i - A_i).$$

La loi d'équilibre des temps d'attentes $\{W_j, j \geq 1\}$ est donnée par $F_w(x) = \mathbb{P}[W_j \leq x] = 1 - \rho e^{-(\mu-\lambda)x}$ pour $x \geq 0$, où $\rho = \lambda/\mu$.

Si W_1 suit cette loi, chaque W_j aussi et on a

$$w = \mathbb{E}[W_j] = \frac{\rho}{\mu(1-\rho)} \quad \text{et} \quad \sigma^2 = \text{Var}[W_j] = \frac{\rho(2-\rho)}{\mu^2(1-\rho)^2}.$$

(Important: ce n'est pas équivalent à démarrer le processus à $W_1 = w$.)

On a aussi une formule pour ρ_j , $\lim_{j \rightarrow \infty} \rho_{j+1}/\rho_j = 4\rho/(1+\rho)^2 < 1$, et

$$\sigma_\infty^2 = \frac{\rho(\rho^3 - 4\rho^2 + 5\rho + 2)}{\mu^2(1-\rho)^4}.$$

Pour ρ proche de 1, $\sigma_\infty^2 \approx 4\rho/[\mu^2(1-\rho)^4] \approx 4\sigma^2/(1-\rho)^2 \gg \sigma^2$.

Par exemple, si $\mu = 1$ et $\rho = 0.9$, alors $\rho_{j+1}/\rho_j \rightarrow 3.6/3.61 \approx 0.9972$, $\rho_{200} \approx 0.30$, $\sigma^2 = 99$ et $\sigma_\infty^2 = 39890$.

Exemple: File $M/M/1$. Taux d'arrivée λ , taux de service μ .

$$W_{i+1} = \max(0, W_i + S_i - A_i).$$

La loi d'équilibre des temps d'attentes $\{W_j, j \geq 1\}$ est donnée par $F_w(x) = \mathbb{P}[W_j \leq x] = 1 - \rho e^{-(\mu-\lambda)x}$ pour $x \geq 0$, où $\rho = \lambda/\mu$.

Si W_1 suit cette loi, chaque W_j aussi et on a

$$w = \mathbb{E}[W_j] = \frac{\rho}{\mu(1-\rho)} \quad \text{et} \quad \sigma^2 = \text{Var}[W_j] = \frac{\rho(2-\rho)}{\mu^2(1-\rho)^2}.$$

(Important: ce n'est pas équivalent à démarrer le processus à $W_1 = w$.)

On a aussi une formule pour ρ_j , $\lim_{j \rightarrow \infty} \rho_{j+1}/\rho_j = 4\rho/(1+\rho)^2 < 1$, et

$$\sigma_\infty^2 = \frac{\rho(\rho^3 - 4\rho^2 + 5\rho + 2)}{\mu^2(1-\rho)^4}.$$

Pour ρ proche de 1, $\sigma_\infty^2 \approx 4\rho/[\mu^2(1-\rho)^4] \approx 4\sigma^2/(1-\rho)^2 \gg \sigma^2$.

Par exemple, si $\mu = 1$ et $\rho = 0.9$, alors $\rho_{j+1}/\rho_j \rightarrow 3.6/3.61 \approx 0.9972$, $\rho_{200} \approx 0.30$, $\sigma^2 = 99$ et $\sigma_\infty^2 = 39890$.

Whitt (1989) donne des approximations de σ_∞^2 pour des quantités autres que W_j ,

Exemple: Une CM irréductible ayant un nombre fini d'états.

Soit $\{Y_j, j \geq 0\}$ une CM avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} , et loi d'équilibre $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$.

Exemple: Une CM irréductible ayant un nombre fini d'états.

Soit $\{Y_j, j \geq 0\}$ une CM avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} , et loi d'équilibre $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$.

On a $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$.

Exemple: Une CM irréductible ayant un nombre fini d'états.

Soit $\{Y_j, j \geq 0\}$ une CM avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} , et loi d'équilibre $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$.

On a $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$. Si $\mathbb{P}[Y_0 = i] = \pi_i$, la CM est stationnaire.

Exemple: Une CM irréductible ayant un nombre fini d'états.

Soit $\{Y_j, j \geq 0\}$ une CM avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} , et loi d'équilibre $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$.

On a $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$. Si $\mathbb{P}[Y_0 = i] = \pi_i$, la CM est stationnaire.

Supposons que chaque visite dans l'état i coûte c_i . Si $\mathbf{c} = (c_1, \dots, c_k)^\text{t}$, on a

$$\mu = \boldsymbol{\pi}\mathbf{c} = \sum_{i=1}^k \pi_i c_i.$$

Exemple: Une CM irréductible ayant un nombre fini d'états.

Soit $\{Y_j, j \geq 0\}$ une CM avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} , et loi d'équilibre $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$.

On a $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$. Si $\mathbb{P}[Y_0 = i] = \pi_i$, la CM est stationnaire.

Supposons que chaque visite dans l'état i coûte c_i . Si $\mathbf{c} = (c_1, \dots, c_k)^\mathbf{t}$, on a

$$\mu = \boldsymbol{\pi}\mathbf{c} = \sum_{i=1}^k \pi_i c_i.$$

Soit \mathbf{D} une matrice diagonale avec éléments π_1, \dots, π_k , et $\mathbf{\Pi}$ une matrice carrée dont toutes les lignes sont $\boldsymbol{\pi}$. On a

$$\begin{aligned} \sigma^2 \rho_j &= \text{Cov}[X_0, X_j] = \mathbf{c}^\mathbf{t} \mathbf{D} (\mathbf{P}^j - \mathbf{\Pi}) \mathbf{c}, \\ \sigma_\infty^2 &= \mathbf{c}^\mathbf{t} \mathbf{D} [\mathbf{I} - \mathbf{\Pi} + 2(\mathbf{I} - \mathbf{P}\mathbf{\Pi})^{-1}(\mathbf{P} - \mathbf{\Pi})] \mathbf{c}. \end{aligned}$$

Exemple: Une CM irréductible ayant un nombre fini d'états.

Soit $\{Y_j, j \geq 0\}$ une CM avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} , et loi d'équilibre $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$.

On a $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$. Si $\mathbb{P}[Y_0 = i] = \pi_i$, la CM est stationnaire.

Supposons que chaque visite dans l'état i coûte c_i . Si $\mathbf{c} = (c_1, \dots, c_k)^\mathbf{t}$, on a

$$\mu = \boldsymbol{\pi}\mathbf{c} = \sum_{i=1}^k \pi_i c_i.$$

Soit \mathbf{D} une matrice diagonale avec éléments π_1, \dots, π_k , et $\mathbf{\Pi}$ une matrice carrée dont toutes les lignes sont $\boldsymbol{\pi}$. On a

$$\begin{aligned} \sigma^2 \rho_j &= \text{Cov}[X_0, X_j] = \mathbf{c}^\mathbf{t} \mathbf{D} (\mathbf{P}^j - \mathbf{\Pi}) \mathbf{c}, \\ \sigma_\infty^2 &= \mathbf{c}^\mathbf{t} \mathbf{D} [\mathbf{I} - \mathbf{\Pi} + 2(\mathbf{I} - \mathbf{P}\mathbf{\Pi})^{-1}(\mathbf{P} - \mathbf{\Pi})] \mathbf{c}. \end{aligned}$$

On a $\rho_j = \mathcal{O}(\beta^j)$ et $\sigma_\infty^2 = \mathcal{O}(1/(1 - \beta))$ où $\beta < 1$ est la deuxième plus grande valeur propre (en valeur absolue) de \mathbf{P} .

Détection et réduction du biais initial

Générer l'état initial selon la loi d'équilibre: aucun biais, mais difficile.

Détection et réduction du biais initial

Générer l'état initial selon la loi d'équilibre: aucun biais, mais difficile.

Techniques basées sur le couplage des trajectoires....

Détection et réduction du biais initial

Générer l'état initial selon la loi d'équilibre: aucun biais, mais difficile.

Techniques basées sur le couplage des trajectoires....

Si $(\mu_i, \sigma_i^2, \rho_{i,i+j}) \rightarrow (\mu, \sigma^2, \rho_j)$ quand $i \rightarrow \infty$, comment détecter et réduire le biais $|\mathbb{E}[\bar{C}_n] - \mu|$ ou $|\mathbb{E}[V_{N(t)}]/t - \bar{v}|$, qui dépend de l'état initial?

Les solutions pratiques sont heuristiques.

La méthode régénérative (à venir) contourne le problème.

Temps discret. Processus $\{C_i, i \geq 1\}$.

Typiquement, $|\mathbb{E}[C_i] - \mu| = \mathcal{O}(\beta^i)$ pour $\beta < 1$. Dans ce cas,

$$|\mathbb{E}[\bar{C}_n] - \mu| = \mathcal{O}\left(\frac{1 + \beta + \beta^2 + \dots}{n}\right) = \mathcal{O}\left(\frac{1}{(1 - \beta)n}\right).$$

Temps discret. Processus $\{C_i, i \geq 1\}$.

Typiquement, $|\mathbb{E}[C_i] - \mu| = \mathcal{O}(\beta^i)$ pour $\beta < 1$. Dans ce cas,

$$|\mathbb{E}[\bar{C}_n] - \mu| = \mathcal{O}\left(\frac{1 + \beta + \beta^2 + \dots}{n}\right) = \mathcal{O}\left(\frac{1}{(1 - \beta)n}\right).$$

Supposons que l'on ne compte pas les n_0 premières observations:

$$\bar{C}_{n_0, n} = \frac{1}{n - n_0} \sum_{j=n_0+1}^n C_j.$$

Temps discret. Processus $\{C_i, i \geq 1\}$.

Typiquement, $|\mathbb{E}[C_i] - \mu| = \mathcal{O}(\beta^i)$ pour $\beta < 1$. Dans ce cas,

$$|\mathbb{E}[\bar{C}_n] - \mu| = \mathcal{O}\left(\frac{1 + \beta + \beta^2 + \dots}{n}\right) = \mathcal{O}\left(\frac{1}{(1 - \beta)n}\right).$$

Supposons que l'on ne compte pas les n_0 premières observations:

$$\bar{C}_{n_0, n} = \frac{1}{n - n_0} \sum_{j=n_0+1}^n C_j.$$

Si $|\mathbb{E}[C_i] - \mu| \leq \kappa_0 \beta^i$ pour $\kappa_0 < \infty$ et $\beta < 1$, alors

$$\begin{aligned} |\mathbb{E}[\bar{C}_{n_0, n}] - \mu| &\leq \frac{\kappa_0}{n - n_0} \sum_{i=n_0+1}^n \beta^i = \frac{\kappa_0 \beta^{n_0+1} (1 - \beta^{n-n_0})}{(n - n_0)(1 - \beta)} \\ &= \mathcal{O}\left(\frac{\beta^{n_0+1}}{(n - n_0)(1 - \beta)}\right) \approx \mathcal{O}\left(\frac{\beta^{n_0}}{n(1 - \beta)}\right) \end{aligned}$$

si $n \gg n_0$.

En augmentant n_0 , on diminue le biais mais on augmente la variance, qui est dans $\mathcal{O}(1/(n - n_0))$.

En augmentant n_0 , on diminue le biais mais on augmente la variance, qui est dans $\mathcal{O}(1/(n - n_0))$.

Pour n fixé, difficile en pratique d'optimiser n_0 pour minimiser le MSE.

En augmentant n_0 , on diminue le biais mais on augmente la variance, qui est dans $\mathcal{O}(1/(n - n_0))$.

Pour n fixé, difficile en pratique d'optimiser n_0 pour minimiser le MSE.
Heuristiques...

En augmentant n_0 , on diminue le biais mais on augmente la variance, qui est dans $\mathcal{O}(1/(n - n_0))$.

Pour n fixé, difficile en pratique d'optimiser n_0 pour minimiser le MSE.
Heuristiques...

Quand $n \rightarrow \infty$, le biais² devient négligeable p.r. à la variance.

Temps continu. Idem.

Pour estimer le coût moyen par unité de temps \bar{v} , horizon tronqué t , échauffement $t_0 < t$, estimateur

$$\frac{V_{N(t)} - V_{N(t_0)}}{t - t_0} = \frac{1}{t - t_0} \sum_{j=N(t_0)+1}^{N(t)} C_j.$$

Temps continu. Idem.

Pour estimer le coût moyen par unité de temps \bar{v} , horizon tronqué t , échauffement $t_0 < t$, estimateur

$$\frac{V_{N(t)} - V_{N(t_0)}}{t - t_0} = \frac{1}{t - t_0} \sum_{j=N(t_0)+1}^{N(t)} C_j.$$

Typiquement (e.g., si $N(t)/t \rightarrow \lambda$, une constante), le biais et la variance se comportent, en fonction de t et t_0 , comme dans le cas discret en fonction de n et n_0 :

$$\text{Biais} = \mathcal{O} \left(\frac{\tilde{\beta}^{t_0}}{t(1 - \tilde{\beta})} \right) \text{ et Variance} = \mathcal{O} \left(\frac{1}{t - t_0} \right).$$

Heuristiques.

Mais comment choisir n_0 (ou t_0) en pratique?

Heuristiques.

Mais comment choisir n_0 (ou t_0) en pratique?

Aucune méthode fiable et universelle.

Heuristiques.

Mais comment choisir n_0 (ou t_0) en pratique?

Aucune méthode fiable et universelle.

Heuristique de Welch.

1. Faire k simulations de longueur n_1 .

C_{ij} = observation j de la répétition i .

2. Poser $\bar{C}_j = \sum_{i=1}^k C_{ij}/k$.

3. Lissage des hautes fréquences par une moyenne mobile de largeur w :

$$\bar{C}_j(w) = \frac{1}{2w+1} \sum_{s=-w}^w \bar{C}_{j+s}, \quad \text{pour } j = w+1, \dots, n_1 - w.$$

4. Regarder le graphique de $\bar{C}_j(w)$ en fonction de j et soit n_0 la valeur de j où le processus semble stable (choix **subjectif**).

Horizon tronqué: Une seule ou plusieurs répétitions?

Une fois n_0 (ou t_0) choisi, on fera, disons, k répétitions indépendantes de longueurs T_1, \dots, T_k , où les T_i sont déterministes ou aléatoires.

Horizon tronqué: Une seule ou plusieurs répétitions?

Une fois n_0 (ou t_0) choisi, on fera, disons, k répétitions indépendantes de longueurs T_1, \dots, T_k , où les T_i sont déterministes ou aléatoires.

Pour la répétition i , on pose

$$X_i = \frac{1}{T_i - t_0} \sum_{j=N(t_0)+1}^{N(T_i)} C_j.$$

Estimateur global: $\bar{X}_k = (1/k) \sum_{i=1}^k X_i$.

Horizon tronqué: Une seule ou plusieurs répétitions?

Une fois n_0 (ou t_0) choisi, on fera, disons, k répétitions indépendantes de longueurs T_1, \dots, T_k , où les T_i sont déterministes ou aléatoires.

Pour la répétition i , on pose

$$X_i = \frac{1}{T_i - t_0} \sum_{j=N(t_0)+1}^{N(T_i)} C_j.$$

Estimateur global: $\bar{X}_k = (1/k) \sum_{i=1}^k X_i$.

Si $k > 2$, on peut exploiter le fait que les X_i sont i.i.d. pour estimer la variance et calculer un IC.

Horizon tronqué: Une seule ou plusieurs répétitions?

Une fois n_0 (ou t_0) choisi, on fera, disons, k répétitions indépendantes de longueurs T_1, \dots, T_k , où les T_i sont déterministes ou aléatoires.

Pour la répétition i , on pose

$$X_i = \frac{1}{T_i - t_0} \sum_{j=N(t_0)+1}^{N(T_i)} C_j.$$

Estimateur global: $\bar{X}_k = (1/k) \sum_{i=1}^k X_i$.

Si $k > 2$, on peut exploiter le fait que les X_i sont i.i.d. pour estimer la variance et calculer un IC.

C'est l'approche “replication-deletion.”

Cas discret:

$$X_i = \frac{1}{n - n_0} \sum_{j=n_0+1}^n C_j.$$

Cas discret:

$$X_i = \frac{1}{n - n_0} \sum_{j=n_0+1}^n C_j.$$

Pour un budget de calcul fixe kn , comment choisir k ? Typiquement, $k = 1$ minimise le MSE (e.g., devoir 1), mais pas toujours.

Cas discret:

$$X_i = \frac{1}{n - n_0} \sum_{j=n_0+1}^n C_j.$$

Pour un budget de calcul fixe kn , comment choisir k ? Typiquement, $k = 1$ minimise le MSE (e.g., devoir 1), mais pas toujours.

Plus difficile d'estimer la variance lorsque $k = 1$ (à voir).

Cas discret:

$$X_i = \frac{1}{n - n_0} \sum_{j=n_0+1}^n C_j.$$

Pour un budget de calcul fixe kn , comment choisir k ? Typiquement, $k = 1$ minimise le MSE (e.g., devoir 1), mais pas toujours.

Plus difficile d'estimer la variance lorsque $k = 1$ (à voir).

Le MSE est minimisé pour $k > 1$ lorsque

- (a) les autocorrélations diminuent très lentement et
- (b) le biais initial est très faible ou diminue très vite.

Cas discret:

$$X_i = \frac{1}{n - n_0} \sum_{j=n_0+1}^n C_j.$$

Pour un budget de calcul fixe kn , comment choisir k ? Typiquement, $k = 1$ minimise le MSE (e.g., devoir 1), mais pas toujours.

Plus difficile d'estimer la variance lorsque $k = 1$ (à voir).

Le MSE est minimisé pour $k > 1$ lorsque

- (a) les autocorrélations diminuent très lentement et
- (b) le biais initial est très faible ou diminue très vite.

Exemple pire cas: voir notes.

Exemple. Supposons que $|\mathbb{E}[C_j] - \mu| = \kappa_0 \beta^j$ et $\text{Cov}[C_i, C_{i+j}] = \sigma^2 \rho_j = \sigma^2 \alpha^j$ pour $j \geq 0$, où $\beta < 1$, $0 < \alpha < 1$, $\kappa_0 > 0$, et $\sigma^2 > 0$.

Exemple. Supposons que $|\mathbb{E}[C_j] - \mu| = \kappa_0 \beta^j$ et $\text{Cov}[C_i, C_{i+j}] = \sigma^2 \rho_j = \sigma^2 \alpha^j$ pour $j \geq 0$, où $\beta < 1$, $0 < \alpha < 1$, $\kappa_0 > 0$, et $\sigma^2 > 0$.

Budget de calcul $B = nk$. On veut choisir k et n_0 pour minimiser

$$\begin{aligned} \text{MSE}[\bar{X}_k] &= (\mathbb{E}[X_i - \mu])^2 + \text{Var}[X_i]/k \\ &= \left(\frac{\kappa_0(\beta^{n_0+1} - \beta^{n+1})}{(n - n_0)(1 - \beta)} \right)^2 + \frac{\sigma^2}{(n - n_0)k} \left(1 + 2 \sum_{j=1}^{n-n_0-1} \frac{\alpha^j (n - n_0 - j)}{n - n_0} \right). \end{aligned}$$

Exemple. Supposons que $|\mathbb{E}[C_j] - \mu| = \kappa_0 \beta^j$ et $\text{Cov}[C_i, C_{i+j}] = \sigma^2 \rho_j = \sigma^2 \alpha^j$ pour $j \geq 0$, où $\beta < 1$, $0 < \alpha < 1$, $\kappa_0 > 0$, et $\sigma^2 > 0$.

Budget de calcul $B = nk$. On veut choisir k et n_0 pour minimiser

$$\begin{aligned} \text{MSE}[\bar{X}_k] &= (\mathbb{E}[X_i - \mu])^2 + \text{Var}[X_i]/k \\ &= \left(\frac{\kappa_0(\beta^{n_0+1} - \beta^{n+1})}{(n - n_0)(1 - \beta)} \right)^2 + \frac{\sigma^2}{(n - n_0)k} \left(1 + 2 \sum_{j=1}^{n-n_0-1} \frac{\alpha^j (n - n_0 - j)}{n - n_0} \right). \end{aligned}$$

Il faut prendre $k > 1$ si β et κ_0 sont petits et si σ^2 et α sont grands.

Si $n \rightarrow \infty$ pour n_0 fixé, le biais $\mathbb{E}[X_i] - \mu$ tend vers zero et le TLCF nous assure que X_1, \dots, X_k sont i.i.d. normales à la limite.

Si $n \rightarrow \infty$ pour n_0 fixé, le biais $\mathbb{E}[X_i] - \mu$ tend vers zero et le TLCHF nous assure que X_1, \dots, X_k sont i.i.d. normales à la limite.

Mais si n_0 et n sont fixés et $k \rightarrow \infty$, le biais ne disparaît pas et le TLC devient

$$\frac{\sqrt{k}(\bar{X}_k - \mu)}{\sigma_{n_0, n}} \Rightarrow N(0, 1) + \frac{\sqrt{k}\beta_{n_0, n}}{\sigma_{n_0, n}}$$

où $\beta_{n_0, n} = \mathbb{E}[X_i] - \mu$ et $\sigma_{n_0, n}^2 = \text{Var}[X_i]$.

Si $n \rightarrow \infty$ pour n_0 fixé, le biais $\mathbb{E}[X_i] - \mu$ tend vers zero et le TLCHF nous assure que X_1, \dots, X_k sont i.i.d. normales à la limite.

Mais si n_0 et n sont fixés et $k \rightarrow \infty$, le biais ne disparaît pas et le TLC devient

$$\frac{\sqrt{k}(\bar{X}_k - \mu)}{\sigma_{n_0, n}} \Rightarrow N(0, 1) + \frac{\sqrt{k}\beta_{n_0, n}}{\sigma_{n_0, n}}$$

où $\beta_{n_0, n} = \mathbb{E}[X_i] - \mu$ et $\sigma_{n_0, n}^2 = \text{Var}[X_i]$.

Un IC basé sur l'hypothèse que les X_i sont i.i.d. normales est asymptotiquement valide seulement si $\sqrt{k}\beta_{n_0, n}/\sigma_{n_0, n} \rightarrow 0$ quand $k \rightarrow \infty$.

Si $n \rightarrow \infty$ pour n_0 fixé, le biais $\mathbb{E}[X_i] - \mu$ tend vers zero et le TLCHF nous assure que X_1, \dots, X_k sont i.i.d. normales à la limite.

Mais si n_0 et n sont fixés et $k \rightarrow \infty$, le biais ne disparaît pas et le TLC devient

$$\frac{\sqrt{k}(\bar{X}_k - \mu)}{\sigma_{n_0, n}} \Rightarrow N(0, 1) + \frac{\sqrt{k}\beta_{n_0, n}}{\sigma_{n_0, n}}$$

où $\beta_{n_0, n} = \mathbb{E}[X_i] - \mu$ et $\sigma_{n_0, n}^2 = \text{Var}[X_i]$.

Un IC basé sur l'hypothèse que les X_i sont i.i.d. normales est asymptotiquement valide seulement si $\sqrt{k}\beta_{n_0, n}/\sigma_{n_0, n} \rightarrow 0$ quand $k \rightarrow \infty$.

Autrement dit, le biais doit converger vers zero plus vite que l'écart-type de \bar{X}_k .

Exemple (suite).

Si $|\mathbb{E}[C_j] - \mu| = \kappa_0 \beta^j$ et $\rho_j = \sigma^2 \alpha^j$, alors

$$\beta_{n_0, n} = \frac{\kappa_0(\beta^{n_0+1} - \beta^{n+1})}{(n - n_0)(1 - \beta)} = \mathcal{O}\left(\frac{\beta^{n_0}}{n - n_0}\right)$$

et

$$\frac{\sigma_{n_0, n}^2}{k} = \frac{\sigma^2}{(n - n_0)k} \left(1 + 2 \sum_{j=1}^{n - n_0 - 1} \frac{\alpha^j (n - n_0 - j)}{n - n_0} \right) = \mathcal{O}\left(\frac{1}{(n - n_0)k}\right).$$

Exemple (suite).

Si $|\mathbb{E}[C_j] - \mu| = \kappa_0 \beta^j$ et $\rho_j = \sigma^2 \alpha^j$, alors

$$\beta_{n_0, n} = \frac{\kappa_0(\beta^{n_0+1} - \beta^{n+1})}{(n - n_0)(1 - \beta)} = \mathcal{O}\left(\frac{\beta^{n_0}}{n - n_0}\right)$$

et

$$\frac{\sigma_{n_0, n}^2}{k} = \frac{\sigma^2}{(n - n_0)k} \left(1 + 2 \sum_{j=1}^{n-n_0-1} \frac{\alpha^j (n - n_0 - j)}{n - n_0} \right) = \mathcal{O}\left(\frac{1}{(n - n_0)k}\right).$$

Ainsi,

$$\frac{\sqrt{k} \beta_{n_0, n}}{\sigma_{n_0, n}} = \mathcal{O}\left(\sqrt{k/(n - n_0)} \beta^{n_0}\right).$$

$\rightarrow 0$ lorsque $k \rightarrow \infty$ ssi

$$\frac{\ln[k/(n - n_0)]}{2 \ln \beta} + n_0 \rightarrow \infty.$$

Exemple (suite).

Si $|\mathbb{E}[C_j] - \mu| = \kappa_0 \beta^j$ et $\rho_j = \sigma^2 \alpha^j$, alors

$$\beta_{n_0, n} = \frac{\kappa_0(\beta^{n_0+1} - \beta^{n+1})}{(n - n_0)(1 - \beta)} = \mathcal{O}\left(\frac{\beta^{n_0}}{n - n_0}\right)$$

et

$$\frac{\sigma_{n_0, n}^2}{k} = \frac{\sigma^2}{(n - n_0)k} \left(1 + 2 \sum_{j=1}^{n-n_0-1} \frac{\alpha^j (n - n_0 - j)}{n - n_0} \right) = \mathcal{O}\left(\frac{1}{(n - n_0)k}\right).$$

Ainsi,

$$\frac{\sqrt{k} \beta_{n_0, n}}{\sigma_{n_0, n}} = \mathcal{O}\left(\sqrt{k/(n - n_0)} \beta^{n_0}\right).$$

$\rightarrow 0$ lorsque $k \rightarrow \infty$ ssi

$$\frac{\ln[k/(n - n_0)]}{2 \ln \beta} + n_0 \rightarrow \infty.$$

Si on augmente k , il faut aussi augmenter n et n_0 assez vite.

IC basé sur une seule longue simulation

Supposons ici que $\{C_j, j \geq 1\}$ est stationnaire, avec $\mathbb{E}[C_j] = \mu$.
(E.g., la partie “échauffement” a déjà été enlevée.)

On estime μ par \bar{C}_n . Comment estimer $\text{Var}[\bar{C}_n]$?

Analyse spectrale

Une première idée serait simplement d'estimer les autocorrélations ρ_k dans l'expression

$$\text{Var}[\bar{C}_n] = \frac{\sigma^2}{n} \left(1 + \frac{2}{n} \sum_{k=1}^{n-1} (n-k) \rho_k \right).$$

Analyse spectrale

Une première idée serait simplement d'estimer les autocorrélations ρ_k dans l'expression

$$\text{Var}[\bar{C}_n] = \frac{\sigma^2}{n} \left(1 + \frac{2}{n} \sum_{k=1}^{n-1} (n-k) \rho_k \right).$$

Difficultés: beaucoup de bruit et les estimateurs standards sont biaisés lorsque k est grand.

Moyennes par lots (“batch means”)

C'est l'approche la plus simple et la plus populaire pour les systèmes complexes.

Moyennes par lots (“batch means”)

C'est l'approche la plus simple et la plus populaire pour les systèmes complexes.

Idée: Regrouper les n observations en k lots de taille $\ell = n/k$.

Moyennes par lots (“batch means”)

C’est l’approche la plus simple et la plus populaire pour les systèmes complexes.

Idée: Regrouper les n observations en k lots de taille $\ell = n/k$.

Moyenne pour le lot i :

$$X_i = \frac{1}{\ell} \sum_{j=\ell(i-1)+1}^{\ell i} C_j.$$

Moyenne globale:

$$\bar{X}_k = \frac{1}{k} \sum_{i=1}^k X_i = \bar{C}_n.$$

Moyennes par lots (“batch means”)

C’est l’approche la plus simple et la plus populaire pour les systèmes complexes.

Idée: Regrouper les n observations en k lots de taille $\ell = n/k$.

Moyenne pour le lot i :

$$X_i = \frac{1}{\ell} \sum_{j=\ell(i-1)+1}^{\ell i} C_j.$$

Moyenne globale:

$$\bar{X}_k = \frac{1}{k} \sum_{i=1}^k X_i = \bar{C}_n.$$

Même principe pour le cas du temps continu (voir notes).

Moyennes par lots (“batch means”)

C'est l'approche la plus simple et la plus populaire pour les systèmes complexes.

Idée: Regrouper les n observations en k lots de taille $\ell = n/k$.

Moyenne pour le lot i :

$$X_i = \frac{1}{\ell} \sum_{j=\ell(i-1)+1}^{\ell i} C_j.$$

Moyenne globale:

$$\bar{X}_k = \frac{1}{k} \sum_{i=1}^k X_i = \bar{C}_n.$$

Même principe pour le cas du **temps continu** (voir notes).

Si ℓ est grand, on s'attend à ce que les X_i soient très peu corrélés et suivent à peu près la loi normale.

Moyennes par lots (“batch means”)

C’est l’approche la plus simple et la plus populaire pour les systèmes complexes.

Idée: Regrouper les n observations en k lots de taille $\ell = n/k$.

Moyenne pour le lot i :

$$X_i = \frac{1}{\ell} \sum_{j=\ell(i-1)+1}^{\ell i} C_j.$$

Moyenne globale:

$$\bar{X}_k = \frac{1}{k} \sum_{i=1}^k X_i = \bar{C}_n.$$

Même principe pour le cas du **temps continu** (voir notes).

Si ℓ est grand, on s’attend à ce que les X_i soient très peu corrélés et suivent à peu près la loi normale.

On peut alors calculer un IC en faisant l’hypothèse que les X_i sont i.i.d. $N(\mu, \sigma_x^2)$, et que donc $\sqrt{k}(\bar{X}_k - \mu)/S_k \sim \text{Student}(k - 1)$.

Il est souvent recommandé de choisir ℓ le plus grand possible et $k \leq 30$.

Il est souvent recommandé de choisir ℓ le plus grand possible et $k \leq 30$.

Un k plus petit diminue le **bias** $\mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k]$, mais augmente **Var** $[S_k^2/k]$.

Il est souvent recommandé de choisir ℓ le plus grand possible et $k \leq 30$.

Un k plus petit diminue le **bias** $\mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k]$, mais augmente **Var** $[S_k^2/k]$.

Que dit une analyse théorique plus minutieuse?

Il est souvent recommandé de choisir ℓ le plus grand possible et $k \leq 30$.

Un k plus petit diminue le **bias** $\mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k]$, mais augmente **Var** $[S_k^2/k]$.

Que dit une analyse théorique plus minutieuse?

Dans le cas où notre **TLCF** tient, on peut montrer que

$$\begin{aligned} \text{Cov}[X_i, X_{i+1}] &= \mathcal{O}(1/\ell), \\ \mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k] &= \mathcal{O}(1/(\ell n)), \quad (\text{biais}) \\ \text{Var}[S_k^2/k] &= \mathcal{O}(\ell/n^3). \quad (\text{variance}) \end{aligned}$$

Les X_i sont aussi asymptotiquement i.i.d. normaux quand $\ell \rightarrow \infty$ et la méthode est justifiée.

Il est souvent recommandé de choisir ℓ le plus grand possible et $k \leq 30$.

Un k plus petit diminue le **bias** $\mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k]$, mais augmente **Var** $[S_k^2/k]$.

Que dit une analyse théorique plus minutieuse?

Dans le cas où notre **TLCF** tient, on peut montrer que

$$\begin{aligned}\text{Cov}[X_i, X_{i+1}] &= \mathcal{O}(1/\ell), \\ \mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k] &= \mathcal{O}(1/(\ell n)), \quad (\text{biais}) \\ \text{Var}[S_k^2/k] &= \mathcal{O}(\ell/n^3). \quad (\text{variance})\end{aligned}$$

Les X_i sont aussi asymptotiquement i.i.d. normaux quand $\ell \rightarrow \infty$ et la méthode est justifiée. De plus,

$$\text{MSE}[S_k^2/k] = \mathcal{O}(1/(\ell n)^2 + \ell/n^3)$$

et pour optimiser le taux de convergence de cette borne quand $n \rightarrow \infty$, il faut prendre

Il est souvent recommandé de choisir ℓ le plus grand possible et $k \leq 30$.

Un k plus petit diminue le **bias** $\mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k]$, mais augmente **Var** $[S_k^2/k]$.

Que dit une analyse théorique plus minutieuse?

Dans le cas où notre **TLCF** tient, on peut montrer que

$$\begin{aligned}\text{Cov}[X_i, X_{i+1}] &= \mathcal{O}(1/\ell), \\ \mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k] &= \mathcal{O}(1/(\ell n)), \quad (\text{biais}) \\ \text{Var}[S_k^2/k] &= \mathcal{O}(\ell/n^3). \quad (\text{variance})\end{aligned}$$

Les X_i sont aussi asymptotiquement i.i.d. normaux quand $\ell \rightarrow \infty$ et la méthode est justifiée. De plus,

$$\text{MSE}[S_k^2/k] = \mathcal{O}(1/(\ell n)^2 + \ell/n^3)$$

et pour optimiser le taux de convergence de cette borne quand $n \rightarrow \infty$, il faut prendre $\ell = \mathcal{O}(n^{1/3})$ et donc $k = \mathcal{O}(n^{2/3})$.

Il est souvent recommandé de choisir ℓ le plus grand possible et $k \leq 30$.

Un k plus petit diminue le **bias** $\mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k]$, mais augmente **Var** $[S_k^2/k]$.

Que dit une analyse théorique plus minutieuse?

Dans le cas où notre **TLCF** tient, on peut montrer que

$$\begin{aligned}\text{Cov}[X_i, X_{i+1}] &= \mathcal{O}(1/\ell), \\ \mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k] &= \mathcal{O}(1/(\ell n)), \quad (\text{biais}) \\ \text{Var}[S_k^2/k] &= \mathcal{O}(\ell/n^3). \quad (\text{variance})\end{aligned}$$

Les X_i sont aussi asymptotiquement i.i.d. normaux quand $\ell \rightarrow \infty$ et la méthode est justifiée. De plus,

$$\text{MSE}[S_k^2/k] = \mathcal{O}(1/(\ell n)^2 + \ell/n^3)$$

et pour optimiser le taux de convergence de cette borne quand $n \rightarrow \infty$, il faut prendre $\ell = \mathcal{O}(n^{1/3})$ et donc $k = \mathcal{O}(n^{2/3})$.

Pour n fixé, le ℓ optimal dépend des constantes cachées.

Il est souvent recommandé de choisir ℓ le plus grand possible et $k \leq 30$.

Un k plus petit diminue le **bias** $\mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k]$, mais augmente **Var** $[S_k^2/k]$.

Que dit une analyse théorique plus minutieuse?

Dans le cas où notre **TLCF** tient, on peut montrer que

$$\begin{aligned}\text{Cov}[X_i, X_{i+1}] &= \mathcal{O}(1/\ell), \\ \mathbb{E}[S_k^2/k] - \text{Var}[\bar{X}_k] &= \mathcal{O}(1/(\ell n)), \quad (\text{biais}) \\ \text{Var}[S_k^2/k] &= \mathcal{O}(\ell/n^3). \quad (\text{variance})\end{aligned}$$

Les X_i sont aussi asymptotiquement i.i.d. normaux quand $\ell \rightarrow \infty$ et la méthode est justifiée. De plus,

$$\text{MSE}[S_k^2/k] = \mathcal{O}(1/(\ell n)^2 + \ell/n^3)$$

et pour optimiser le taux de convergence de cette borne quand $n \rightarrow \infty$, il faut prendre $\ell = \mathcal{O}(n^{1/3})$ et donc $k = \mathcal{O}(n^{2/3})$.

Pour n fixé, le ℓ optimal dépend des constantes cachées.

En principe, on peut l'estimer.

Il existe des logiciels (gratuits) qui implantent des méthodes qui déterminent n et k de manière **adaptative**, jusqu'à ce que la précision voulue soit atteinte.

LABATCH.2 (Fishman 1998);

ASAP2 (Steiner et al. 2002).

Il existe des logiciels (gratuits) qui implantent des méthodes qui déterminent n et k de manière **adaptative**, jusqu'à ce que la précision voulue soit atteinte.

LABATCH.2 (Fishman 1998);

ASAP2 (Steiner et al. 2002).

ASAP2 n'essaie pas de rendre les X_i indépendants, mais augmente ℓ jusqu'à ce qu'ils soient approx. normaux, puis ajustent un modèle **AR(1)** à la suite des X_i pour pouvoir estimer $\text{Var}[\bar{X}_k]$.

Simulation régénérative

Idée: écrire la moyenne sur horizon infini comme un rapport de deux espérances sur horizon fini.

Simulation régénérative

Idée: écrire la moyenne sur horizon infini comme un rapport de deux espérances sur horizon fini.

Définition. Un processus stochastique $\{Y(t), t \geq 0\}$ est **régénératif** (au sens classique) s'il existe une variable aléatoire $\tau_1 > 0$ telle que $\{Y(\tau_1 + t), t \geq 0\}$ est stochastiquement équivalent à $\{Y(t), t \geq 0\}$ et indépendant de τ_1 et de $\{Y(t), t < \tau_1\}$.

Simulation régénérative

Idée: écrire la moyenne sur horizon infini comme un rapport de deux espérances sur horizon fini.

Définition. Un processus stochastique $\{Y(t), t \geq 0\}$ est **régénératif** (au sens classique) s'il existe une variable aléatoire $\tau_1 > 0$ telle que $\{Y(\tau_1 + t), t \geq 0\}$ est stochastiquement équivalent à $\{Y(t), t \geq 0\}$ et indépendant de τ_1 et de $\{Y(t), t < \tau_1\}$.

La v.a. τ_1 est un **instant de régénération**. La trajectoire du processus sur l'intervalle de temps $(0, \tau_1]$ est un **cycle régénératif**.

Simulation régénérative

Idée: écrire la moyenne sur horizon infini comme un rapport de deux espérances sur horizon fini.

Définition. Un processus stochastique $\{Y(t), t \geq 0\}$ est **régénératif** (au sens classique) s'il existe une variable aléatoire $\tau_1 > 0$ telle que $\{Y(\tau_1 + t), t \geq 0\}$ est stochastiquement équivalent à $\{Y(t), t \geq 0\}$ et indépendant de τ_1 et de $\{Y(t), t < \tau_1\}$.

La v.a. τ_1 est un **instant de régénération**. La trajectoire du processus sur l'intervalle de temps $(0, \tau_1]$ est un **cycle régénératif**.

L'adaptation au cas discret est évidente.

Simulation régénérative

Idée: écrire la moyenne sur horizon infini comme un rapport de deux espérances sur horizon fini.

Définition. Un processus stochastique $\{Y(t), t \geq 0\}$ est **régénératif** (au sens classique) s'il existe une variable aléatoire $\tau_1 > 0$ telle que $\{Y(\tau_1 + t), t \geq 0\}$ est stochastiquement équivalent à $\{Y(t), t \geq 0\}$ et indépendant de τ_1 et de $\{Y(t), t < \tau_1\}$.

La v.a. τ_1 est un **instant de régénération**. La trajectoire du processus sur l'intervalle de temps $(0, \tau_1]$ est un **cycle régénératif**.

L'adaptation au cas discret est évidente.

Si $\{Y(t), t \geq 0\}$ est régénératif, alors $\{Y(\tau_1 + t), t \geq 0\}$ est aussi régénératif avec un instant de régénération τ_2 , etc.

Simulation régénérative

Idée: écrire la moyenne sur horizon infini comme un rapport de deux espérances sur horizon fini.

Définition. Un processus stochastique $\{Y(t), t \geq 0\}$ est **régénératif** (au sens classique) s'il existe une variable aléatoire $\tau_1 > 0$ telle que $\{Y(\tau_1 + t), t \geq 0\}$ est stochastiquement équivalent à $\{Y(t), t \geq 0\}$ et indépendant de τ_1 et de $\{Y(t), t < \tau_1\}$.

La v.a. τ_1 est un **instant de régénération**. La trajectoire du processus sur l'intervalle de temps $(0, \tau_1]$ est un **cycle régénératif**.

L'adaptation au cas discret est évidente.

Si $\{Y(t), t \geq 0\}$ est régénératif, alors $\{Y(\tau_1 + t), t \geq 0\}$ est aussi régénératif avec un instant de régénération τ_2 , etc.

On a ainsi une suite infinie d'instants de régénération $0 = \tau_0 < \tau_1 < \tau_2 \dots$ et de cycles régénératifs i.i.d..

Simulation régénérative

Idée: écrire la moyenne sur horizon infini comme un rapport de deux espérances sur horizon fini.

Définition. Un processus stochastique $\{Y(t), t \geq 0\}$ est **régénératif** (au sens classique) s'il existe une variable aléatoire $\tau_1 > 0$ telle que $\{Y(\tau_1 + t), t \geq 0\}$ est stochastiquement équivalent à $\{Y(t), t \geq 0\}$ et indépendant de τ_1 et de $\{Y(t), t < \tau_1\}$.

La v.a. τ_1 est un **instant de régénération**. La trajectoire du processus sur l'intervalle de temps $(0, \tau_1]$ est un **cycle régénératif**.

L'adaptation au cas discret est évidente.

Si $\{Y(t), t \geq 0\}$ est régénératif, alors $\{Y(\tau_1 + t), t \geq 0\}$ est aussi régénératif avec un instant de régénération τ_2 , etc.

On a ainsi une suite infinie d'instant de régénération $0 = \tau_0 < \tau_1 < \tau_2 \dots$ et de cycles régénératifs i.i.d..

Ces τ_j ne sont pas nécessairement des temps d'arrêt.

Si $\mathbb{E}[\tau_1] < \infty$, le processus est dit **récurrent positif**.

Si $\mathbb{E}[\tau_1] < \infty$, le processus est dit **récurrent positif**.

Parfois, le premier cycle est différent des autres: le processus est régénératif **avec délai** (à partir de τ_1). Effet négligeable à long terme.

Si $\mathbb{E}[\tau_1] < \infty$, le processus est dit **récurrent positif**.

Parfois, le premier cycle est différent des autres: le processus est régénératif **avec délai** (à partir de τ_1). Effet négligeable à long terme.

Si $\{Y(t), t \geq 0\}$ est régénératif aux instants τ_1, τ_2, \dots et si $C(t) = f(Y(t))$, alors $\{C(t), t \geq 0\}$ est aussi régénératif aux mêmes instants.

Example. Soit $\{Y_i, i \geq 0\}$ une CM en temps discret avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} telle que tous les états communiquent, et $Y_0 = j$.

Example. Soit $\{Y_i, i \geq 0\}$ une CM en temps discret avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} telle que tous les états communiquent, et $Y_0 = j$.

Ce processus est régénératif, avec instants de régénération aux époques i où $Y_i = j$. Il est aussi récurrent positif.

Example. Soit $\{Y_i, i \geq 0\}$ une CM en temps discret avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} telle que tous les états communiquent, et $Y_0 = j$.

Ce processus est régénératif, avec instants de régénération aux époques i où $Y_i = j$. Il est aussi récurrent positif.

Si $Y_0 = k \neq j$, le processus est régénératif avec délai.

Exemple. Soit $\{Y_i, i \geq 0\}$ une CM en temps discret avec espace d'états fini $\{1, \dots, k\}$, matrice de transition \mathbf{P} telle que tous les états communiquent, et $Y_0 = j$.

Ce processus est régénératif, avec instants de régénération aux époques i où $Y_i = j$. Il est aussi récurrent positif.

Si $Y_0 = k \neq j$, le processus est régénératif avec délai.

En général, les points de régénération ne correspondent pas nécessairement à la visite d'un état particulier. Par exemple, si toutes les lignes de \mathbf{P} sont identiques, on peut poser $\tau_i = i$ pour tout i .

Example. File $GI/GI/1$ stable: $W_{i+1} = \max(0, W_i + S_i - A_i)$.

Si $W_1 = 0$, si les v.a. $S_i - A_i$ sont i.i.d. et $\mathbb{E}[S_i - A_i] < 0$, alors $\{W_i, i \geq 1\}$ est régénératif et on peut prendre pour instants de régénération les époques i où $W_i = 0$.

Example. File $GI/GI/1$ stable: $W_{i+1} = \max(0, W_i + S_i - A_i)$.

Si $W_1 = 0$, si les v.a. $S_i - A_i$ sont i.i.d. et $\mathbb{E}[S_i - A_i] < 0$, alors $\{W_i, i \geq 1\}$ est régénératif et on peut prendre pour instants de régénération les époques i où $W_i = 0$.

Le processus $\{Q(t), t \geq 0\}$ régénère aussi, aux instants où un client arrive dans un système vide.

Example. File $GI/GI/1$ stable: $W_{i+1} = \max(0, W_i + S_i - A_i)$.

Si $W_1 = 0$, si les v.a. $S_i - A_i$ sont i.i.d. et $\mathbb{E}[S_i - A_i] < 0$, alors $\{W_i, i \geq 1\}$ est régénératif et on peut prendre pour instants de régénération les époques i où $W_i = 0$.

Le processus $\{Q(t), t \geq 0\}$ régénère aussi, aux instants où un client arrive dans un système vide.

Peut-on prendre les instants où le système se vide comme instants de régénération? Seulement dans le cas $M/G/1$.

Example. Centre d'appels.

Supposons qu'il opère pour une suite infinie de jours i.i.d..
Il ouvre à 8h et ferme à 21h.

Example. Centre d'appels.

Supposons qu'il opère pour une suite infinie de jours i.i.d..

Il ouvre à 8h et ferme à 21h.

Soit t le temps écoulé depuis le premier jour à minuit, en heures.

$Q(t)$ = nombre d'appels dans la file au temps t ;

$\{Q(t), t \geq 0\}$ régénère à $\tau_j = 24j$ pour $j = 1, 2, \dots$

Example. Centre d'appels.

Supposons qu'il opère pour une suite infinie de jours i.i.d..

Il ouvre à 8h et ferme à 21h.

Soit t le temps écoulé depuis le premier jour à minuit, en heures.

$Q(t)$ = nombre d'appels dans la file au temps t ;

$\{Q(t), t \geq 0\}$ régénère à $\tau_j = 24j$ pour $j = 1, 2, \dots$

Si X_j = nombre d'appels reçus au jour j , $\{X_j, j \geq 0\}$ régénère à $\tau_j = j$ pour tout j (**processus de renouvellement**).

Example. Centre d'appels.

Supposons qu'il opère pour une suite infinie de jours i.i.d..

Il ouvre à 8h et ferme à 21h.

Soit t le temps écoulé depuis le premier jour à minuit, en heures.

$Q(t)$ = nombre d'appels dans la file au temps t ;

$\{Q(t), t \geq 0\}$ régénère à $\tau_j = 24j$ pour $j = 1, 2, \dots$

Si X_j = nombre d'appels reçus au jour j , $\{X_j, j \geq 0\}$ régénère à $\tau_j = j$ pour tout j (**processus de renouvellement**).

Même chose pour $\{Z_j, j \geq 0\}$ si Z_j = nombre d'abandons au jour j .

Théorème du renouvellement avec gains

Soit $\{C_i, i \geq 0\}$ régénératif aux instants $\tau_j, j \geq 0$.

Théorème du renouvellement avec gains

Soit $\{C_i, i \geq 0\}$ régénératif aux instants $\tau_j, j \geq 0$. Posons

$$V_n = \sum_{j=1}^n C_j \quad (\text{coût total pour les } n \text{ premiers événements}),$$

$$X_j = V_{N(\tau_j)} - V_{N(\tau_{j-1})} \quad (\text{coût pour le cycle } j),$$

$$Y_j = \tau_j - \tau_{j-1} \quad (\text{durée du cycle } j).$$

Théorème du renouvellement avec gains

Soit $\{C_i, i \geq 0\}$ régénératif aux instants $\tau_j, j \geq 0$. Posons

$$V_n = \sum_{j=1}^n C_j \quad (\text{coût total pour les } n \text{ premiers événements}),$$

$$X_j = V_{N(\tau_j)} - V_{N(\tau_{j-1})} \quad (\text{coût pour le cycle } j),$$

$$Y_j = \tau_j - \tau_{j-1} \quad (\text{durée du cycle } j).$$

Théorème. Si $\mathbb{E}[Y_j] > 0$ et $\mathbb{E}[|X_j|] < \infty$, alors,

$$\bar{v} \stackrel{\text{def}}{=} \lim_{t \rightarrow \infty} \frac{\mathbb{E}[V_{N(t)}]}{t} = \frac{\mathbb{E}[X_j]}{\mathbb{E}[Y_j]} \quad (\text{version espérance}),$$

et

$$\bar{v} \stackrel{\text{p.s.}}{=} \lim_{t \rightarrow \infty} \frac{V_{N(t)}}{t} \quad (\text{version trajectoire}).$$

Intervalles de confiance

Si on fixe le nombre de cycles n , on se ramène au problème d'estimer le quotient $\bar{v} = \mathbb{E}[X_j]/\mathbb{E}[Y_j]$ à partir des observations i.i.d. $(X_1, Y_1), \dots, (X_n, Y_n)$. Connue.

Intervalles de confiance

Si on fixe le nombre de cycles n , on se ramène au problème d'estimer le quotient $\bar{v} = \mathbb{E}[X_j]/\mathbb{E}[Y_j]$ à partir des observations i.i.d. $(X_1, Y_1), \dots, (X_n, Y_n)$. Connus.

Si on fixe la durée de la simulation à t , le nombre de cycles

$M(t) = \sup\{n \geq 0 : \tau_n \leq t\}$ est aléatoire.

Intervalles de confiance

Si on fixe le nombre de cycles n , on se ramène au problème d'estimer le quotient $\bar{v} = \mathbb{E}[X_j]/\mathbb{E}[Y_j]$ à partir des observations i.i.d. $(X_1, Y_1), \dots, (X_n, Y_n)$. Connus.

Si on fixe la durée de la simulation à t , le nombre de cycles

$M(t) = \sup\{n \geq 0 : \tau_n \leq t\}$ est aléatoire.

Lorsqu'on atteint t , on peut ou bien

- (a) terminer le cycle en cours (on en aura $M(t) + 1$);
- (b) laisser tomber le cycle en cours (on en aura $M(t)$).

Intervalles de confiance

Si on fixe le nombre de cycles n , on se ramène au problème d'estimer le quotient $\bar{v} = \mathbb{E}[X_j]/\mathbb{E}[Y_j]$ à partir des observations i.i.d. $(X_1, Y_1), \dots, (X_n, Y_n)$. Connus.

Si on fixe la durée de la simulation à t , le nombre de cycles $M(t) = \sup\{n \geq 0 : \tau_n \leq t\}$ est aléatoire.

Lorsqu'on atteint t , on peut ou bien

- (a) terminer le cycle en cours (on en aura $M(t) + 1$);
- (b) laisser tomber le cycle en cours (on en aura $M(t)$).

La variance est dans $\mathcal{O}(1/t)$ dans les deux cas.

Le biais sur \bar{v} est dans $\mathcal{O}(1/t^2)$ pour (a) et $\mathcal{O}(1/t)$ pour b.

Intervalles de confiance

Si on fixe le nombre de cycles n , on se ramène au problème d'estimer le quotient $\bar{v} = \mathbb{E}[X_j]/\mathbb{E}[Y_j]$ à partir des observations i.i.d. $(X_1, Y_1), \dots, (X_n, Y_n)$. Connue.

Si on fixe la durée de la simulation à t , le nombre de cycles

$M(t) = \sup\{n \geq 0 : \tau_n \leq t\}$ est aléatoire.

Lorsqu'on atteint t , on peut ou bien

- (a) terminer le cycle en cours (on en aura $M(t) + 1$);
- (b) laisser tomber le cycle en cours (on en aura $M(t)$).

La variance est dans $\mathcal{O}(1/t)$ dans les deux cas.

Le biais sur \bar{v} est dans $\mathcal{O}(1/t^2)$ pour (a) et $\mathcal{O}(1/t)$ pour b.

Théorème. Sous les conditions du TLC pour un quotient, quand $t \rightarrow \infty$,

$$\frac{\sqrt{M(t)}(\hat{\mu}_{M(t)} - \mu)}{\hat{\sigma}_{z,M(t)}} \Rightarrow \frac{\sqrt{t/\bar{Y}_{M(t)}}(\hat{\mu}_{M(t)} - \mu)}{\hat{\sigma}_{z,M(t)}} \Rightarrow \frac{\sqrt{t/\mathbb{E}[Y_1]}(\hat{\mu}_{M(t)} - \mu)}{\hat{\sigma}_{z,M(t)}} \Rightarrow N(0, 1).$$

Valide aussi si on remplace $M(t)$ par $M(t) + 1$.

Coûts (ou revenus) actualisés

Taux d'actualisation $\rho > 0$.

Pour un état initial \mathcal{S}_0 , on veut estimer

$$v_\rho^\infty = \sum_{i=1}^{\infty} \mathbb{E}[e^{-\rho t_i} C_i] = \mathbb{E}[V_\rho^\infty]$$

(on suppose) où $\{C_i, i \geq 0\}$ est régénératif et

$$\begin{aligned} V_\rho^\infty &= \sum_{i=1}^{\infty} e^{-\rho t_i} C_i \\ &= V_{\rho, N(\tau_1)} + \sum_{i=N(\tau_1)+1}^{\infty} e^{-\rho t_i} C_i \\ &= V_{\rho, N(\tau_1)} + e^{-\rho \tau_1} \sum_{i=N(\tau_1)+1}^{\infty} e^{-\rho(t_i - \tau_1)} C_i. \end{aligned}$$

$$V_{\rho}^{\infty} = V_{\rho, N(\tau_1)} + e^{-\rho\tau_1} \sum_{i=N(\tau_1)+1}^{\infty} e^{-\rho(t_i-\tau_1)} C_i.$$

En prenant l'espérance:

$$\mathbb{E}[V_{\rho}^{\infty}] = \mathbb{E}[V_{\rho, N(\tau_1)}] + \mathbb{E}[e^{-\rho\tau_1}] \mathbb{E}[V_{\rho}^{\infty}]$$

$$V_{\rho}^{\infty} = V_{\rho, N(\tau_1)} + e^{-\rho\tau_1} \sum_{i=N(\tau_1)+1}^{\infty} e^{-\rho(t_i-\tau_1)} C_i.$$

En prenant l'espérance:

$$\mathbb{E}[V_{\rho}^{\infty}] = \mathbb{E}[V_{\rho, N(\tau_1)}] + \mathbb{E}[e^{-\rho\tau_1}] \mathbb{E}[V_{\rho}^{\infty}]$$

et donc

$$v_{\rho}^{\infty} = \mathbb{E}[V_{\rho}^{\infty}] = \frac{\mathbb{E}[V_{\rho, N(\tau_1)}]}{1 - \mathbb{E}[e^{-\rho\tau_1}]}.$$

On s'est ramené à estimer un quotient de deux espérances sur horizon fini.