

# IFT 3245

## Simulation et modèles

Fabian Bastin  
DIRO  
Université de Montréal

Automne 2012

# Intervalle basé sur une simulation unique

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

Si nous estimons  $\mu$  par  $\bar{C}_n$ , comment pouvons-nous estimer  $\text{Var}[\bar{C}_n]$ ?

Deux techniques particulières:

- moyennes par lots;
- simulation régénérative.

# Moyennes par lots (“batch means”)

C'est l'approche la plus simple et la plus populaire pour les systèmes complexes. L'idée de base consiste à regrouper les  $n$  observations en  $k$  lots de taille  $\ell = n/k$ . Nous pouvons alors écrire

- 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”)

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

Sous certaines conditions additionnelles, il est possible de prouver que c'est le cas pour  $\ell \rightarrow \infty$ , et  $k$  fixé. Nous pouvons alors calculer un intervalle de confiance 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  $\ell$  le plus grand possible et  $k \leq 30$ . Un  $k$  plus petit diminue le bias  $E[S_k^2/k] - \text{Var}[\bar{X}_k]$ , mais augmente la variance de  $S_k^2/k$ .

L'idée consiste à écrire la moyenne sur horizon infini comme un rapport de deux espérances sur horizon fini.

## Définition

*Processus régénératif* 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  $\tau_1$  et de  $\{Y(t), t < \tau_1\}$ .

La variable aléatoire  $\tau_1$  est un instant de régénération. La trajectoire du processus sur l'intervalle de temps  $(0, \tau_1]$  est un cycles régénératif, et si  $E[\tau_1] < \infty$ , le processus est dit récurrent positif.

L'adaptation au cas discret est évidente.

# Simulation régénérative

Si  $\{Y(t), t \geq 0\}$  est régénératif, alors  $\{Y(\tau_1 + t), t \geq 0\}$  est aussi régénératif avec un instant de régénération  $\tau_2$ , 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.

Parfois, le premier cycle est différent des autres: le processus est régénératif avec délai (à partir de  $\tau_1$ ), mais l'effet est négligeable à long terme.

Si  $\{Y(t), t \geq 0\}$  est régénératif aux instants  $\tau_1, \tau_2, \dots$  et si  $C(t) = f(Y(t))$ , alors  $\{C(t), t \geq 0\}$  est aussi régénératif aux mêmes instants.

# Simulation régénérative: exemple

File  $GI/G/1$  stable.

Si  $W_1 = 0$ , si les v.a.  $S_i - A_i$  sont i.i.d. et  $E[S_i - A_i] < 0$ , 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.

Mais peut-on prendre les instants où le système se vide comme instants de régénération? La réponse est positive seulement dans le cas  $M/G/1$ .

# Simulation régénérative: exemple

## Centre d'appels

Supposons qu'un centre d'appels 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, et  $Q(t) =$  le nombre d'appels dans la file au temps  $t$ . Le processus  $\{Q(t), t \geq 0\}$  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).

Il en est de même pour  $\{Z_j, j \geq 0\}$  si  $Z_j =$  nombre d'abandons au jour  $j$ .

# Théorème du renouvellement avec gains

Soit  $\{C_i, i \geq 0\}$  un processus régénératif aux instants  $\tau_j, j \geq 0$ , et  $N(t)$  le nombre d'événements qui se sont produits au cours de l'intervalle  $[0, t]$ .

Posons

$$X_j = V_{N(\tau_j)} - V_{N(\tau_{j-1})},$$

le coût pour le cycle  $j$ , et

$$Y_j = \tau_j - \tau_{j-1},$$

la durée du cycle  $j$ .

# Théorème du renouvellement avec gains

## Theorème

Si  $E[Y_j] > 0$  et  $E[|X_j|] < \infty$ , alors,

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

et

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

# Théorème du renouvellement avec gains

Une fois le nombre de cycles  $n$  fixé, nous retrouvons le problème d'estimation du quotient  $\bar{v} = E[X_j]/E[Y_j]$  à partir des observations i.i.d.  $(X_1, Y_1), \dots, (X_n, Y_n)$ .

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

Lorsque nous atteignons  $t$ , nous pouvons terminer le cycle en cours (nous aura  $M(t) + 1$  cycles), ou ignorer le cycle en cours (nous en aurons  $M(t)$ ). La variance est dans  $O(1/t)$  dans les deux cas. Le biais sur  $\bar{v}$  est dans  $O(1/t^2)$  pour la première approche, et  $O(1/t)$  pour la seconde.

# Théorème du renouvellement avec gains

## Theorème

*Sous les conditions du théorème de la limite centrale pour un quotient, quand  $t \rightarrow \infty$ ,*

$$\begin{aligned}\frac{\sqrt{M(t)}(\hat{v}_{M(t)} - v)}{\hat{\sigma}_{g,M(t)}} &\xrightarrow{D} \frac{\sqrt{t/\bar{Y}_{M(t)}}(\hat{v}_{M(t)} - v)}{\hat{\sigma}_{g,M(t)}} \\ &\xrightarrow{D} \frac{\sqrt{t/E[Y_1]}(\hat{v}_{M(t)} - v)}{\hat{\sigma}_{g,M(t)}} \\ &\xrightarrow{D} N(0, 1),\end{aligned}$$

où  $\sigma_g$  est défini comme dans le cadre du théorème delta.

Le résultat demeure valide si nous remplaçons  $M(t)$  par  $M(t) + 1$ .