Analyse de sortie

Plan

- Estimateur et Intervalle de confiance
- Sortie de simulation
- Méthodes de réduction de variance

Mesures de performance

Considérer l'estimation d'un paramètre de performance θ (ou φ), d'une simulation

- Données temporelles discrètes:

$$[Y_1, Y_2, \dots, Y_n]$$

avec la moyenne θ.

- Données temporelles continues:

$$\{Y(t), 0 \le t \le T_E\}$$

avec moyenne pondérée dans le temps ϕ.

Estimateur

Pour les données temporelles discrètes.

- L'estimateur ponctuel:

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

  - est sans biaisé si sa espérance est θ,

$$E\,[\theta] = \theta$$

  - est biaisé si

$$E\,[\theta] \neq \theta$$

Estimateur

Pour les données en temps continu.

- L'estimateur ponctuel:

$$\hat{\theta} = \frac{1}{T_E}\int_0^{T_E} Y(t)dt$$

  - est biaisé en général où

$$E\,[\hat{\theta}] = \theta$$

  - est biaisé si

$$E\,[\theta] \neq \theta$$

Estimateur

Mesures compatibles:

- Taux d'occupation d'un serveur

$$Y(t) = \begin{cases} 1 & si\ nb\ de\ clients\ servis > 0 \\ 0 & si\ nb\ de\ clients\ servis = 0 \end{cases}$$

Mesures non compatibles :

- Quantiles
- Distributions

Estimateur

Considérons $X_i \; i = 1, \dots, n$

- Estimateur de $\mu = E(X)$ sans biais

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$Var(\bar{X}) = Var\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{1}{n^2}Var[X_1 + \cdots + X_n]$$

$$= \frac{1}{n^2}Var[X_1] + \frac{1}{n^2}Var[X_n] = \frac{Var[X]}{n} = \frac{\sigma^2}{n}$$

Intervalle de confiance

- Inégalité de Chebychev ($n \to \infty$)

$$P\{|\bar{X} - \mu| \geq \varepsilon\} \leq \frac{\sigma^2/n}{\varepsilon^2}$$

- Théorème central limite ($n > 30$)

$$P\{|\bar{X} - \mu| \geq \varepsilon\} \approx P\left\{|Z| > \frac{\varepsilon\sqrt{n}}{\sigma}\right\} = 2\left[1 - \Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right)\right]$$

où Φ est la fonction de distribution de N(0,1).

Intervalle de confiance

- Avec $Var(X)$ connue

$$[\bar{X} - \frac{z_{1-\alpha/2}\sigma}{\sqrt{n}}, \bar{X} + \frac{z_{1-\alpha/2}\sigma}{\sqrt{n}}]$$

- Avec $Var(X)$ inconnue

$$[\bar{X} - \frac{z_{1-\alpha/2}S}{\sqrt{n}}, \bar{X} + \frac{z_{1-\alpha/2}S}{\sqrt{n}}]$$

avec

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

Intervalle de confiance

- Loi de Student (n petit)

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim T^{n-1}$$

suit une loi de Student avec degré de liberté $n - 1$

- Intervalle de confiance

$$\left[\bar{X} - t_{\alpha/2}^{n-1}\frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2}^{n-1}\frac{S}{\sqrt{n}}\right]$$

Type de simulations

- Simulation transitoire ou avec terminal
- Simulation stationnaire ou sans terminal
- Le fait qu'une simulation soit considérée comme terminée ou non dépend de:
  - Les objectifs de l'étude de simulation et
  - La nature du système.

Type de simulations

Simulation transitoire :

- Fonctionne pendant une certaine durée T_E, où E est un événement spécifié qui arrête la simulation.
- Commence à l'heure 0 dans des conditions initiales bien définies.
- Se termine au temps d'arrêt T_E.

Simulation transitoire

Simulation transitoire : Exemple bancaire:

- Ouverture à 8h30 (heure 0) sans présence de clients et 8 des 11 guichets travaillant (conditions initiales), et fermeture à 16h30 (heure T_E = 480 minutes).
- L'analyste de simulation choisit de le considérer comme une simulation transitoire car l'objet d'intérêt est l'opération d'un jour.

Simulation transitoire

Méthode de répétitions indépendantes

- Les observations indépendantes ou échantillons peuvent être obtenus en exécutant la même simulation plusieurs fois avec exactement les mêmes conditions initiales et le même critère d'arrêt.
- La durée d'exécution souvent courte => grand nombre de répétitions possible.

Simulation transitoire

Soient

- N le nombre de simulations exécutées
- n le nombre d'observations obtenues de chaque itération
- X_{ik} le k-ème échantillon obtenue dans la i-ème simulation

Alors l'estimateur final de X

$$\bar{\bar{X}} = \frac{1}{N}\sum_{i=1}^{N} \bar{X}_i$$

avec $\bar{X}_i = \frac{1}{n}\sum_{k=1}^{n} X_{ik}$ indépendents.

Simulation d'état stationnaire

Simulation d'état stationnaire ou sans terminal

- Fonctionne en continu, ou au moins sur une très longue période de temps.
- Exemples:
  - lignes d'assemblage qui s'arrêtent rarement,
  - systèmes téléphoniques,
  - salles d'urgence des hôpitaux.

Simulation d'état stationnaire

Simulation d'état stationnaire ou sans terminal

- Conditions initiales définies par l'analyste.
- Fonctionne pendant une certaine période de temps définie par l'analyste T_E.
- Étudier les propriétés en régime permanent (à long terme) du système, propriétés qui ne sont pas influencées par les conditions initiales du modèle.

Problème : longue durée de simulation avant d'atteindre à l'état stationnaire

=> La simulation s'arrête si un certain phase (une période) transitoire a été détecté

Simulation d'état stationnaire

Méthode de la moyenne par lots

- Seulement une simulation lancée=> une période « warmup »
- Enlever la phase transitoire pour cette exécution
- Couper la simulation en N simulations partielles ou lots

La corrélation entre X_i et X_j est

- forte si i n'est pas loin de j
- Faible si i est loin de j

Simulation d’état stationnaire

- Si la largueur des lots est suffisamment grande, les estimateurs dérivés de chaque lots sont approximativement indépendants.
- Soit L est le nombre tel que la corrélation entre X_i et X_{i+L} est négligeable.
- La longueur de lot doit $\geq 5L$
- Au moins 20 – 30 lots

Simulation d’état stationnaire

N : le nombre de lots

- Moyenne empirique de X du i-ème lot

$$\overline{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij}$$

- Moyenne globale

$$\overline{\overline{X}} = \frac{1}{N}\sum_{i=1}^{N} \overline{X}_i$$

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(\overline{X}_i - \overline{\overline{X}}\right)^2$$

Simulation d’état stationnaire

- Autocorrélation (décalage temporel=1)

$$\tilde{\rho}_1 = \frac{\sum_{i=1}^{N}(\overline{X}_i - \overline{\overline{X}})(\overline{X_{i+1}} - \overline{\overline{X}})}{\sum_{i=1}^{N}\left(\overline{X}_i - \overline{\overline{X}}\right)^2}$$

- Si $\tilde{\rho}_1$ petit => lots indépendants

Méthode de réduction de variance

- Échantillonnage préférentiel
- Variable de contrôle
- Variables antithétiques
- Stratification
- …..

Variable de contrôle

Obtenir un nouveau ensemble d’échantillons

$Y_i, i = 1, \ldots, n$ tel que

- identiquement distribués
- Corrélés avec X_i
- $E(Y_i) = v$ connue

Objectifs :

- Vérification de $\bar{X}$ avec $\bar{Y}$
- Réduction de variance de $\bar{X}$

Variable de contrôle

- Un nouveau estimateur de $E(X) = \mu$

$$\overline{X_c} = \bar{X} - c(\bar{Y} - v)$$

  - $\bar{X}$: estimateur de X sans bias

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

  - $\bar{Y}$: estimateur de Y sans bias

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

Variable de contrôle

- Espérance de $\overline{X_c}$

$$
\begin{aligned}
E(\overline{X_c}) &= E\big(\bar{X} - c(\bar{Y} - v)\big) \\
&= E(\bar{X}) - E\big(c(\bar{Y} - v)\big) \\
&= E(\bar{X}) - c\big(E(\bar{Y}) - E(v)\big) \\
&= E(\bar{X}) - c(v - v) \\
&= E(\bar{X}) - 0 \\
&= E(X) \\
&= \mu
\end{aligned}
$$

Variable de contrôle

- Variance de $\overline{X_c}$

$$
\begin{aligned}
Var(\overline{X_c}) &= \mathrm{Var}[\bar{X} - c(\bar{Y} - v)] \\
&= Var[\bar{X} - c\bar{Y}] \\
&= Var[\bar{X} - c] + c^2 Var(\bar{Y}) - 2c\mathrm{Cov}(\bar{X}, \bar{Y})
\end{aligned}
$$

- $Var(\overline{X_c}) < Var(\bar{X})$ si et seulement si

$$c^2 Var(\bar{Y}) - 2c\mathrm{Cov}(\bar{X}, \bar{Y}) < 0$$

Variable de contrôle

- En minimisant la valeur de cette équation quadratique, on trouve la valeur optimale de c

$$c_{opt} = \frac{\mathrm{Cov}(\bar{X}, \bar{Y})}{Var(\bar{Y})}$$

- Donc

$$
\begin{aligned}
Var(\overline{X_c}) &= Var(\bar{X}) + \left(\frac{\mathrm{Cov}(\bar{X}, \bar{Y})}{Var(\bar{Y})}\right)^2 Var(\bar{Y}) - 2\left(\frac{\mathrm{Cov}(\bar{X}, \bar{Y})}{Var(\bar{Y})}\right)\mathrm{Cov}(\bar{X}, \bar{Y}) \\
&= Var(\bar{X}) - \frac{\big(\mathrm{Cov}(\bar{X}, \bar{Y})\big)^2}{Var(\bar{Y})}
\end{aligned}
$$

Variable de contrôle

Estimateurs :

$$S_{X,Y} = \frac{\sum_{i=1}^{n}(X_i - \bar{X}_i)(Y_i - \bar{Y}_i)}{n-1}$$

$$S_Y^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2}{n-1}$$

$$\hat{c}_{opt} = \frac{\sum_{i=1}^{n}(X_i - \bar{X}_i)(Y_i - \bar{Y}_i)}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2}$$

$$S_c^2 = \frac{S_X^2 + S_Y^2 - 2S_{X,Y}}{n}$$