

Discrete event simulation

Fabian Bastin

`fabian.bastin@umontreal.ca`

Université de Montréal – CIRRELT – IVADO – Fin-ML

Basic concepts

System collection d'entités qui agissent et interagissent afin d'accomplir une certaine fin logique.

L'état d'un système est la collection de variables nécessaires pour décrire un système à un instant particulier.

Model description simplifiée d'un système, dans le but d'évaluer sa performance ou l'effet de certaines décisions.

Simulation faire évoluer le modèle d'un système en fournissant les entrées appropriées, puis observer et analyser les résultats.

Two main system types: discrete and continuous. Deux types principaux de systèmes: discrets et continus.

Models

Physical vs mathematical model.

Analytical model, numerical model, model with simulation.

Reference textbook: A. M. Law, Simulation Modeling and Analysis, fifth edition, McGraw-Hill, USA, 2015.

Event approach

Assumption: system in which state variables can only change by a countable number of points in time.

Sequence of events $e_0, e_1, e_2 \dots$, occurring at times

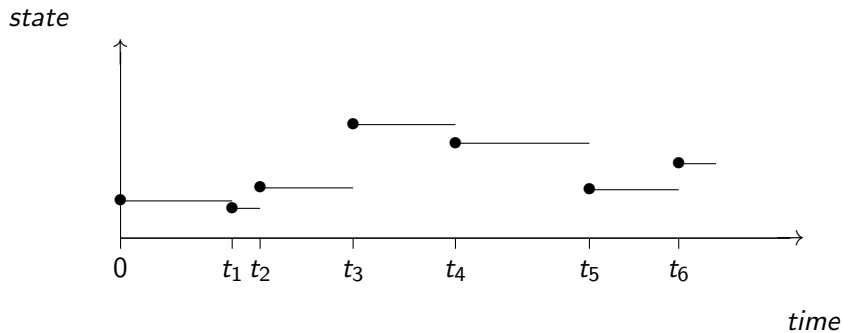
$$0 \leq t_0 \leq t_1 \leq t_2 \leq \dots$$

\mathcal{S}_i : system state immediately after e_i .

simulation time: current value of t_i .

(t_i, \mathcal{S}_i) must contain enough information to continue the simulation (except the realizations of the random variables for events $e_j, j > i$).

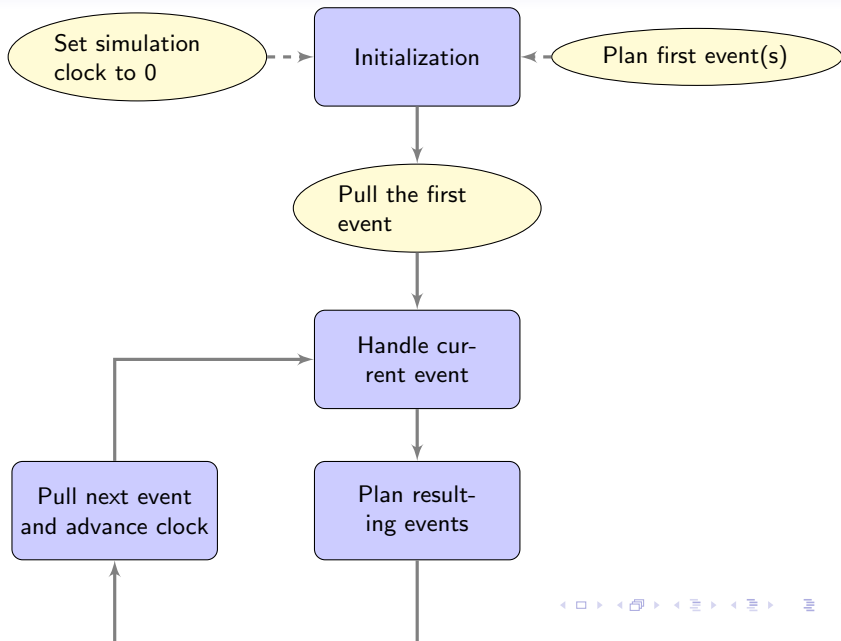
Discrete events

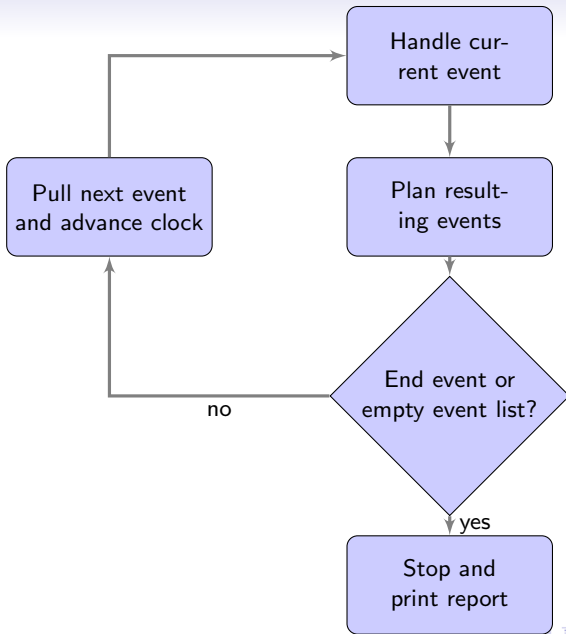


General framework

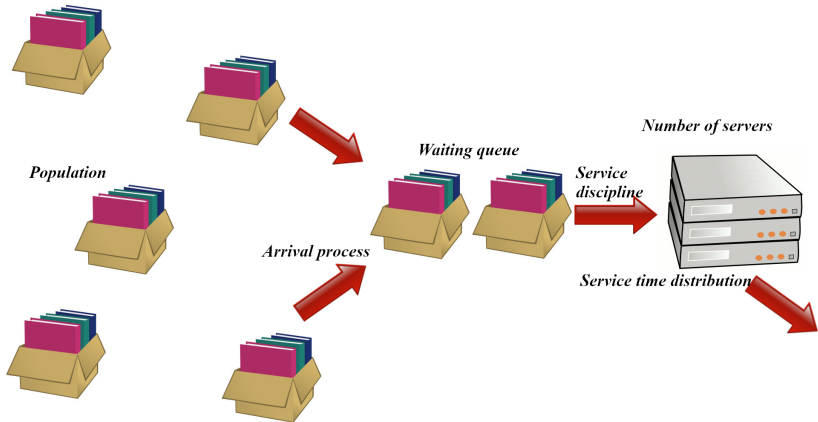
- Main program.
- Simulation clock.
- Event list.
- Random numbers generation functions.
- Statistical collectors.
- Report generation.

Event: action, schedule, cancel.





Queueing systems



Kendall notation

$$A/S/c/K/N/SD$$

- A Arrival process
- S Service time distribution
- c Number of servers
- K Waiting list maximum length
- N Population size
- SD Service discipline

Notation

G General

I Independent

M Markovian or memoryless

Defaults:

- Infinite buffer capacity
- Infinite population size
- FCFS (First-Come-First-Served) service discipline

Memoryless property:

$$P[X > x + a \mid X > a] = P[X > x]$$

Continuous: exponential distribution

Discrete: geometric distribution

Example: GI/G/1 queue

$$G/G/1 = G/G/1/\infty/\infty/FCFS$$

We also write $GI/G/1$ to stress that the service times are independent.

The clients arrive one by one, one being served at a time, under a FCFS priority, with

- S_i : service time of client i , with cdf G ;
- A_i : inter-arrival time between clients i and $i + 1$, with cdf F .

We assume S_i and A_i mutually independent.

Example: GI/G/1 queue

The first client arrives at time A_0 in an empty system.

Nous souhaitons simuler ce système pour une durée T et calculer l'attente moyenne par client et la longueur moyenne de la file d'attente. Les types d'événements sont

1. arrivée;
2. départ;
3. fin de la simulation.

Les variables aléatoires (indépendantes) à générer sont A_1, A_2, A_3, \dots et S_1, S_2, S_3, \dots .

Example: GI/G/1 queue

- W_i , le temps d'attente du client i ;
- $Q(t)$, la longueur de la file d'attente au temps t ;
- $N_c(t)$, le nombre de clients ayant débuté leur service durant l'intervalle $[0, t]$.

Supposons que nous voulons calculer, pour l'intervalle $[0, T]$, l'*attente moyenne* par client,

$$\overline{W}_{N_c(T)} = \frac{1}{N_c(T)} \sum_{i=1}^{N_c(T)} W_i;$$

et la *longueur moyenne* de la file d'attente:

$$\overline{Q}_T = \frac{1}{T} \int_0^T Q(t) dt.$$

Example: GI/G/1 queue

Cette dernière intégrale est facile à calculer par simulation.

Pour chaque *client*, nous créons un *objet* contenant l'instant d'arrivée et la durée de service.

Variables d'état: liste des clients en attente et la liste les clients en service.

+ horloge de simulation, compteurs statistiques (au besoin), *liste* des événements futurs prévus, par ordre chronologique, et une procédure pour chaque type d'événement.

Queueing system simulation

Arrivée Générer A selon F et prévoir une autre arrivée dans A unités de temps;

Créer le nouveau client et noter son instant d'arrivée;

Générer sa durée de service S selon G ;

Si (serveur est occupé) **alors**

Insérer ce client dans la liste des clients en attente;

sinon

Insérer le client dans la liste des clients en service;

Prévoir son départ dans S unités de temps;

Mise à jour des statistiques voulues.

Départ Enlever le client de la liste des clients en service;

Si la file d'attente n'est pas vide **alors**

Enlever le premier client de la liste d'attente;

L'insérer dans la liste des clients en service;

Récupérer son S et prévoir son départ dans S unités de temps;

Mise à jour des statistiques voulues.

Simulation

Démarrer la simulation:

- initialiser des variables et compteurs,
- prévoir la “Fin-de-la-Simulation” au temps T
- générer A selon F et prévoir la première “Arrivée” dans A unités de temps,
- lancement de la simulation.

L'exécution de la simulation consiste simplement à répéter la boucle suivante:

Répéter: exécuter le prochain événement dans la liste d'événements

jusqu'à: la liste des événements prévus est vide, ou bien un événement stoppe la simulation.

Simulation sans liste d'événements.

Pas toujours indispensable de recourir à l'approche par événements.

Exemple: récurrence de Lindley:

$$W_1 = 0, \quad W_{i+1} = \max(0, W_i + S_i - A_{i+1}).$$

Nous pouvons ainsi facilement simuler pour un nombre fixe de clients (au lieu d'un horizon de temps T fixe).

Le formulation du programme basée sur la récurrence de Lindsley revient à traiter un problème d'intégration sur $[0, 1]^t$.

Récurrance de Lindley

Supposons que nous souhaitions estimer $E[\overline{W}_{100}]$ par \overline{W}_{100} :

$$\overline{W}_{100} = \frac{1}{100} \sum_{i=1}^{100} W_i.$$

Il nous faut $A_1, S_1, A_2, S_2, \dots, A_{99}, S_{99}$. Si on pose $A_i = F^{-1}(U_{2i-1})$ et $S_i = G^{-1}(U_{2i})$, où $U_j, j = 1, \dots, 99$, sont des variables aléatoires uniformes sur $(0, 1)$, alors $\overline{W}_{100} = f(U_1, \dots, U_{198})$ (f étant de forme inconnue dans le cas présent).

Nous avons donc $t = 198$, i.e.,

$$E[\overline{W}_{100}] = \int_{[0,1]^{198}} f(\mathbf{u}) d\mathbf{u}.$$

Nombre aléatoire de clients

Par contre, si nous voulons simuler $\overline{W}_{N_c(T)}$, le nombre de clients $N_c(T)$ est aléatoire et non borné. Nous devrions donc choisir $t = \infty$ comme dimension.

Approche par processus

Un processus est une séquence temporellement ordonnée d'événements interreliés, séparés par certains intervalles de temps, qui décrit l'expérience entière d'une "entité" comme celle-ci évolue à travers un "système". Un système ou un modèle de simulation peut avoir différents types de processus.

Une routine sera associée à chaque processus du modèle, décrivant son histoire entière à travers le système.

Au contraire de l'approche par événements, une routine de processus contient explicitement le passage du temps simulé et a généralement de multiples points d'entrée.

Processes vs Events

Une simulation utilisant l'approche par processus évolue aussi au cours du temps en exécutant les événements dans l'ordre de leur occurrence.

En interne, les approches par événement et par processus sont donc très similaires.

Pour l'utilisateur, le raisonnement soit différent, et puisse parfois (mais pas toujours) être plus naturel pour l'approche par processus.

Amélioration de l'Efficacité

The following material is adapted from Pierre L'Ecuyer's class "Simulation: stochastic aspects" ..

L'efficacité d'un estimateur X se définit comme suit:

$$eff[X] = \frac{1}{MSE[X]C(X)},$$

where MSE stands for mean square error and $C(X)$ represents the cost to compute X .

Est-il possible, étant donné X , de construire un nouvel estimateur Y plus efficace que X ?

Nous allons introduire les principaux concepts d'amélioration de l'efficacité au moyen d'un exemple introductif sur les centres d'appels téléphoniques.

Exemple introductif

Posons B , le *facteur d'achalandage* pour la journée, et supposons que $P[B = b_t] = q_t$, où

t	1	2	3	4
b_t	0.8	1.0	1.2	1.4
q_t	0.25	0.55	0.15	0.05

Il est facile de vérifier que $E[B] = 1$.

Les *arrivées* des appels suivent processus de Poisson de taux $B\lambda_j$ durant l'heure j .

Notons $G_i(s)$ = nombre d'appels dont le service a débuté après moins de s secondes d'attente, le jour i .

On veut estimer $\mu = E[G_i(s)]$, disons pour $s = 20$.

Example

Nous supposons de plus que les durées de service des appels suivent la loi $\Gamma(\alpha, \gamma)$, dont la moyenne est $\alpha\gamma$.

Dans notre exemple, on a $\alpha = 1$ et $\gamma = \gamma_1 = 100$.

Nombre d'agents n_j et taux d'arrivée λ_j (par heure) pour 13 périodes d'une heure dans le centre d'appel:

j	0	1	2	3	4	5	6	7	8	9	10	11	12
n_j	4	6	8	8	8	7	8	8	6	6	4	4	4
λ_j	100	150	150	180	200	150	150	150	120	100	80	70	60

We simulate n independent days.

Example

Soit $X_i = G_i(s)$ pour le jour i , et

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

On a $E[\bar{X}_n] = \mu$ et $Var[\bar{X}_n] = Var[X_i]/n$.

Une expérience avec $n = 1000$ donne $\bar{X}_n = 1518.3$ et $S_n^2 = 21615$.
La variance estimée de \bar{X}_n est alors $\widehat{Var}[\bar{X}_n] = 21.6$.

Exemple introductif

Nous souhaitons construire un intervalle de confiance pour $\sigma^2 = n\text{Var}[\bar{X}_n]$, sous l'hypothèse que $(n-1)S_n^2/\sigma^2$ suit approximativement une χ^2 à $n-1$ degrés de liberté. Ceci permet de construire l'intervalle à 90%: $[0.930S_n^2, 1.077S_n^2]$.

En d'autres termes, l'erreur relative de cet estimateur est inférieure à 8% avec une "confiance" d'environ 90%.

Voyons comment *améliorer* cet estimateur \bar{X}_n , en réduisant sa variance. Pour chaque méthode proposée, nous donnerons des résultats numériques pour $n = 1000$.

Estimation indirecte.

Soit A_i le nombre total d'arrivées au jour i ; posons $D_i = A_i - X_i$.

On sait que $a = E[A_i] = \sum_{j=1}^m \lambda_j = 1660$.

Nous pouvons dès lors écrire $\mu = E[X_i] = E[A_i - D_i] = a - E[D_i]$,
que l'on peut estimer par

$$\bar{X}_{i,n} = E[A_i] - \bar{D}_n = a - \frac{1}{n} \sum_{i=1}^n D_i.$$

Cet estimateur a moins de variance que \bar{X}_n ssi $\text{Var}[D_i] < \text{Var}[X_i]$.

Variance estimée: $\widehat{\text{Var}}[X_{i,i}] = 18389$.

Control variate (CV)

Idée: exploiter l'information auxiliaire.

Par exemple, si A_i est plus grand que d'habitude ($A_i > E[A_i] = 1660$), on s'attend à ce que ce jour là, X_i et D_i surestiment $E[X_i]$ et $E[D_i]$.

On pourrait faire une “*correction*” à ces estimateurs: remplacer X_i par

$$X_{c,i} = X_i - \beta(A_i - 1660)$$

où β est une constante appropriée. Alors

$$\bar{X}_{c,n} = \bar{X}_n - \beta(\bar{A}_n - 1660).$$

Control variate (CV)

On a $E[\bar{X}_{c,n}] = E[X_i]$ et

$$\text{Var}[\bar{X}_{c,n}] = \frac{\text{Var}[X_i] + \beta^2 \text{Var}[A_i] - 2\beta \text{Cov}[A_i, X_i]}{n}.$$

Cette variance est une fonction quadratique en β , que l'on minimise en prenant

$$\beta = \beta^* = \text{Cov}[A_i, X_i] / \text{Var}[A_i].$$

La *variance minimale* est

$$\text{Var}[\bar{X}_{c,n}] = \frac{\text{Var}[X_i] - (\beta^*)^2 \text{Var}[A_i]}{n} = \text{Var}[\bar{X}_n](1 - \rho^2[A_i, X_i])$$

où $\rho[A_i, X_i]$ est le coeff. de corrélation entre A_i et X_i .

Control variate (CV)

On ne connaît pas $\text{Cov}[A_i, X_i]$, mais:

- (a) On peut l'estimer par des *expériences pilotes*.
- (b) On peut l'estimer par les *mêmes n simulations* que \bar{X}_n .

Avec (b) on obtient l'estimateur (légèrement biaisé):

$$\bar{X}_{\text{ce},n} = \bar{X}_n - \frac{1}{n-1} \left[\sum_{i=1}^n (A_i - \bar{A}_n)(X_i - \bar{X}_n) \right] \frac{\bar{A}_n - a}{\text{Var}[A_i]}.$$

Conditionnellement à B_i , $A_i \sim \text{Poisson}(1660B_i)$. On a donc

$$\begin{aligned} \text{Var}[A_i] &= \text{Var}[E[A_i|B_i]] + E[\text{Var}[A_i|B_i]] \\ &= \text{Var}[1660B_i] + E[1660B_i] \\ &= 1660^2 \text{Var}[B_i] + 1660E[B_i] \\ &= 67794.4. \end{aligned}$$

Variance empirique obtenue ici: 3310.

Variable de contrôle (VC)

En prenant $\beta = 1$, on retrouve l'estimateur indirect:

$$\bar{X}_{i,n} = \bar{X}_n - (\bar{A}_n - 1660) = 1660 - \bar{D}_n.$$

Si on combine *VC* + *indirect*, on obtient:

$$\begin{aligned}\bar{X}_{i,c,n} &= a - \bar{D}_n - \beta_2(\bar{A}_n - a) \\ &= \bar{A}_n - \bar{D}_n - (1 + \beta_2)(\bar{A}_n - a) \\ &= \bar{X}_n - (1 + \beta_2)(\bar{A}_n - a),\end{aligned}$$

i.e., $\bar{X}_{i,c,n}$ est équivalent à $\bar{X}_{c,n}$ avec $\beta = 1 + \beta_2$. Par conséquent, en présence de la variable de contrôle, l'estimation indirecte n'apporte rien.

Nous pourrions aussi considérer d'autres variables de contrôle, comme B_i , la moyenne des durées de service, etc.

Stratification.

Le facteur d'achalandage B_i est une source importante de variabilité importante dans le cas présent. Essayons de la contrôler.

En posant $\mu_t = E[X_i \mid B_i = b_t]$, on a

$$\begin{aligned}\mu &= E[X_i] = \sum_{t=1}^4 P[B_i = b_t] \cdot E[X_i \mid B_i = b_t] \\ &= .25 E[X_i \mid B_i = 0.8] + .55 E[X_i \mid B_i = 1.0] \\ &\quad + .15 E[X_i \mid B_i = 1.2] + .05 E[X_i \mid B_i = 1.4] \\ &= .25 \mu_1 + .55 \mu_2 + .15 \mu_3 + .05 \mu_4.\end{aligned}$$

Idée: estimer μ_t séparément pour chaque t .

Stratification

Supposons qu'il y a N_t jours où $B_i = b_t$ et soient $X_{t,1}, \dots, X_{t,N_t}$ les valeurs de X_i pour ces jours.

On peut estimer $\mu_t = E[X_i \mid B_i = b_t]$ par

$$\hat{\mu}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} X_{t,i}$$

et μ par

$$\bar{X}_{s,n} = \sum_{t=1}^4 q_t \hat{\mu}_t = .25\hat{\mu}_1 + .55\hat{\mu}_2 + .15\hat{\mu}_3 + .05\hat{\mu}_4.$$

Stratification

On a

$$\begin{aligned} \text{Var}[\bar{X}_{s,n} \mid N_1, N_2, N_3, N_4] \\ &= \sum_{t=1}^4 q_t^2 \text{Var}[\hat{\mu}_t \mid N_t] = \sum_{t=1}^4 q_t^2 \sigma_t^2 / N_t \\ &= .25^2 \sigma_1^2 / N_1 + .55^2 \sigma_2^2 / N_2 + .15^2 \sigma_3^2 / N_3 + .05^2 \sigma_4^2 / N_4. \end{aligned}$$

où $\sigma_t^2 = \text{Var}[X_i \mid B_i = b_t]$.

La variance est réduite si les σ_t^2 sont inférieurs à $\text{Var}[X_i]$.

Si les B_i sont générés normalement: *post-stratification*.

Pour estimer μ par stratification, on peut aussi *fixer les* $N_t = n_t$ à l'avance, c'est-à-dire choisir à l'avance combien de jours on aura $B_i = b_t$ pour chaque valeur de t .

Stratification

- *Allocation proportionnelle*: prendre $n_t = nq_t$.
Avec $n = 1000$, cela donne $n_1 = 250$, $n_2 = 550$, $n_3 = 150$, $n_4 = 50$.
- *Allocation optimale*: choisir les n_t pour minimiser $\text{Var}[\bar{X}_{s,n}]$ sous la contrainte $n_1 + n_2 + n_3 + n_4 = n$. On obtient:

$$\frac{n_t}{n} = \frac{\sigma_t P[B_i = t]}{\sum_{k=1}^4 \sigma_k P[B_i = k]} = \frac{\sigma_k q_k}{\sum_{k=1}^4 \sigma_k q_k}.$$

On ne connaît pas ces σ_k , mais on peut les estimer par des essais pilotes. Avec $n_0 = 800$ essais pilotes, 200 par valeur de t , on obtient par exemple $(n_1, n_2, n_3, n_4) = (219, 512, 182, 87)$ (après arrondi).

Stratégies combinées.

Stratification combinée avec VC:

$$\hat{\mu}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} X_{c,t,i} = \frac{1}{n_t} \sum_{i=1}^{n_t} [X_{t,i} - \beta_t(A_{t,i} - a b_t)].$$

On minimise $\sigma_t^2 = \text{Var}[X_{c,t,i}]$ en prenant

$$\beta_t = \beta_t^* = \text{Cov}[A_{t,i}, X_{t,i}] / \text{Var}[A_{t,i}] = \text{Cov}[A_i, X_i | B_i = b_t] / (a b_t).$$

L'ajout d'une VC change les σ_t^2 : l'allocation optimale n'est plus la même.

Avec $n_0 = 800$ essais pilotes, on obtient

$(\beta_1, \beta_2, \beta_3, \beta_4) = (1.020, 0.648, 0.224, -0.202)$ et

$(n_1, n_2, n_3, n_4) = (131, 503, 247, 119)$ comme estimation des valeurs optimales.

Stratégies combinées.

En répétant l'expérience avec $n = 100000$, on peut trouver les estimations suivantes pour la variance ($\pm 1\%$):

$$\text{Var}[X_n] = 21998; \quad \text{Var}[X_{i,n}] = 17996; \quad \text{Var}[X_{c,n}] = 3043;$$

$$\text{Var}[X_{\text{so},c,n}] = 885.$$

Résultats numériques pour $n = 1000$

Méthode	Estimateur	Mean	$S_n^2(\pm 9\%)$	Ratio
Crude estimator	\bar{X}_n	1518.2	21615	1.000
Indirect	$\bar{X}_{i,n}$	1502.5	18389	0.851
CV A_i , with pilot runs	$\bar{X}_{c,n}$	1510.1	3305	0.153
CV A_i , no pilot runs	$\bar{X}_{ce,n}$	1510.2	3310	0.153
Indirect + CV, no pilot runs	$\bar{X}_{i,c,n}$	1510.1	3309	0.153
Stratification (propor.)	$\bar{X}_{sp,n}$	1509.5	1778	0.082
Stratification (optimal)	$\bar{X}_{so,n}$	1509.4	1568	0.073
Strat. (propor.) + CV	$\bar{X}_{sp,c,n}$	1509.2	1140	0.053
Strat. (optimal) + CV	$\bar{X}_{so,c,n}$	1508.3	900	0.042

Comparaison de deux systèmes similaires

Supposons à présent qu'il est possible de diminuer légèrement le paramètre γ à de la loi gouvernant les durées de service, en prenant

$$\gamma = \gamma_2 = \gamma_1(1 - \delta),$$

pour $\delta > 0$ très petit.

Si on utilise les mêmes nombres aléatoires, cela équivaut à multiplier les durées de service par $(1 - \delta)$.

Nous voulons estimer $\mu(\gamma_2) - \mu(\gamma_1) = E_\gamma[X_2] - E_\gamma[X_1]$, afin de mesurer l'effet d'augmenter un peu la vitesse des serveurs.

Comparaison de deux systèmes similaires

On simule n jours pour chaque valeur de γ .

- $X_{1,i}$ = valeur de $G_i(s)$ avec γ_1 ;
- $X_{2,i}$ = valeur de $G_i(s)$ avec γ_2 ;
- $\Delta_i = X_{2,i} - X_{1,i}$,

$$\bar{\Delta}_n = \frac{1}{n} \sum_{i=1}^n \Delta_i$$

On peut simuler $X_{1,i}$ et $X_{2,i}$

- (i) avec des v.a. indépendantes (VAI),
- (ii) avec des v.a. communes (VAC).

Comparaison de deux systèmes similaires

Comme

$$\text{Var}[\Delta_i] = \text{Var}[X_{1,i}] + \text{Var}[X_{2,i}] - 2\text{Cov}[X_{1,i}, X_{2,i}],$$

le but des variables aléatoires communes est de rendre la covariance positive.

Comment implanter les VAC?

Utiliser des suites aléatoires différentes pour générer:

1. le facteur d'achalandage B_i ;
2. les temps inter-arrivées;
3. les durées des appels;
4. les durées de patience.

Tout est généré par inversion: on utilise une uniforme par v.a.

Synchronisation

Lorsqu'on change les durées de service, les durées d'attente changent et les décisions d'abandon peuvent ainsi changer.

Si on ne génère les durées de service que pour les clients qui n'abandonnent pas, alors on peut perdre la synchronisation: on peut avoir une durée de service de moins à générer dans un système que dans l'autre.

On peut générer les durées de service:

- (a) pour tous les appels, même les abandons,
- (b) seulement pour les appels servis.

Synchronisation

De même, la durée de patience n'a pas besoin d'être générée pour les clients qui n'attendent pas. On peut la générer:

- (c) pour tous les appels,
- (d) seulement si nécessaire.

Synchronization: résultats (avec $n = 10^4$)

Method	$\delta = 0.1$		$\delta = 0.01$		$\delta = 0.001$	
	$\bar{\Delta}_n$	$\widehat{Var}[\Delta_i]$	$\bar{\Delta}_n$	$\widehat{Var}[\Delta_i]$	$\bar{\Delta}_n$	$\widehat{Var}[\Delta_i]$
IRN (a + c)	55.2	56913	4.98	45164	0.66	44046
IRN (a + d)	52.2	54696	7.22	45192	-1.82	45022
IRN (b + c)	50.3	56919	9.98	44241	1.50	45383
IRN (b + d)	53.7	55222	5.82	44659	1.36	44493
CRN, no sync. (b + d)	56.0	3187	5.90	1204	0.19	726
CRN (a + c)	56.4	2154	6.29	37	0.62	1.8
CRN (a + d)	55.9	2161	6.08	158	0.74	53.8
CRN (b + c)	55.8	2333	6.25	104	0.63	7.9
CRN (b + d)	55.5	2323	6.44	143	0.59	35.3

Induction de corrélation

Conditions suffisantes pour que $\text{Cov}[X, Y]$ soit positive, ou soit négative?

Comment maximiser ou minimiser la corrélation pour des lois marginales données?

Theorem (Fréchet bounds)

Parmi les paires de v.a. (X, Y) dont les f.r. marginales sont F et G , la paire $(X, Y) = (F^{-1}(U), G^{-1}(U))$ où $U \sim U(0, 1)$, maximise $\rho[X, Y]$, et la paire $(X, Y) = (F^{-1}(U), G^{-1}(1 - U))$ minimise $\rho[X, Y]$.

Correlation induction

Par exemple, si la fonction de répartition d'une durée de service est F dans le premier système et G dans le second, et si on génère les durées de service par $X = F^{-1}(U)$ et $Y = G^{-1}(U)$, alors $\text{Cov}[X, Y] \geq 0$.

Theorem

Soient $X = f(U)$ où $U \sim U(0, 1)$ et $Y = g(V)$ où $V \sim U(0, 1)$.
Alors,

- si f et g sont monotones dans le même sens et $V = U$, alors $\text{Cov}[X, Y] \geq 0$;
- si f et g sont monotones dans le même sens et $V = 1 - U$, alors $\text{Cov}[X, Y] \leq 0$.

Common random numbers (CRN)

We use $\Delta = X_2 - X_1$ to estimate $\mu_2 - \mu_1 = E[X_2] - E[X_1]$. Then

$$\text{Var}[\Delta] = \text{Var}[X_1] + \text{Var}[X_2] - 2 \text{Cov}[X_1, X_2].$$

Goal: induire une corrélation positive entre X_1 et X_2 sans changer leurs lois individuelles.

Technique: utiliser les mêmes nombres aléatoires pour simuler les deux systèmes, en essayant de maintenir la *synchronisation* le mieux possible.

Si $X_k = f_k(\mathbf{U}_k) = f_k(U_{k,1}, U_{k,2}, \dots)$ pour $k = 1, 2$, utiliser des VAC partout veut dire prendre $\mathbf{U}_1 = \mathbf{U}_2$.

Common random numbers (CRN)

Theorem

Si f_1 et f_2 sont monotones dans le même sens par rapport à tous leurs arguments, alors $\text{Cov}[X_1, X_2] \geq 0$.

Pour *maximiser* la corrélation, il faudrait générer X_1 et X_2 directement par inversion!

Example: Lindley process

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

Si X_1 et X_2 sont des fonctions non-décroissantes des W_i pour deux processus de Lindley simulés avec des VAC, alors $\text{Cov}[X_1, X_2] \geq 0$.

Parfois il est très difficile de vérifier les conditions du théorème.

Par exemple, pour la banque ou un centre d'appels, la monotonie est difficile à vérifier à cause des possibilités d'abandon. De plus, les conditions du théorème sont *suffisantes*, mais *pas nécessaires*.

Common random numbers (CRN)

Pour *tester* si c'est efficace en pratique: faire une *expérience pilote* avec les VAC et estimer $\text{Cov}(X_1, X_2)$.

Il n'est pas nécessaire de faire de simulations sans les VAC: pour estimer la variance qu'on aurait sans les VAC, il suffit de prendre la version empirique de $\text{Var}[X_1] + \text{Var}[X_2]$.