La file M/GI/1

Le processus de Poisson joue un rôle fondamental dans la modélisation d'arrivée de clients dans des files d'attente. L'objectif de ce chapitre est d'étudier une file d'attente avec un seul serveur et des temps de service indépendants identiquement distribués de loi générale. On notera F la fonction de répartition de cette loi et, lorsqu'elle est bien définie f sa fonction de densité. On supposera que la discipline de service est FIFO.

Le processus d'arrivée de clients, noté A_t, est un processus de Poisson d'intensité λ. On se donne comme variable d'état le couple (X_t, R_t) où X_t est le nombre de clients présents dans la file à l'instant t et D_t le temps de service déjà effectué pour le client en service (age du client). Le processus (X_t, R_t) est un processus de Markov à valeur dans $\mathbb{N} \times \mathbb{R}^+$. L'espace d'état est continu, il est par conséquent difficile à analyser directement. On sait cependant que la condition de stabilité du système est

$$\rho \stackrel{\Delta}{=} \lambda \mathbb{E} S < 1,$$

ρ est appelé charge de la file d'attente, correspondant au taux d'utilisation du serveur.

Pour simplifier l'analyse, on étudie le processus à des instants spécifiés (les instants de fin de service). Notons $0 < t_1 < \cdots < t_n < \cdots$ les instants de départ des clients $1, 2, \cdots, n, \cdots$. Le processus $\{X_n\}_{n\in\mathbb{N}}$ défini par

$$X_n = X_{t_n^+},$$

nombre de clients dans la file juste après le départ du $n^{\text{ième}}$ client, est un processus de Markov en temps discret et à valeur dans $\mathbb{N}$. En effet, les équations d'évolution s'écrivent

$$X_{n+1} = \begin{cases} X_n + Y_{n+1} - 1 & \text{si } X_n \geq 1, \\ Y_{n+1} & \text{si } X_n = 0, \end{cases}$$

où $Y_{n+1} = A_{t_{n+1}} - A_{t_n}$ est le nombre de clients arrivés pendant le $(n+1)^{\text{ième}}$ service. Comme les intervalles de temps $]t_i, t_{i+1}]$ sont disjoints et que le processus d'arrivée est un processus de Poisson, les variables aléatoires Y_n sont indépendantes. Comme les temps de service sont indépendants de même loi, les variables aléatoires Y_n sont iid. $\{X_n\}_{n\in\mathbb{N}}$ est donc une chaîne de Markov homogène..

Les probabilités de transition de cette chaîne sont définies par la distribution des variables Y_n. On note

$$a_j = \mathbb{P}(Y_n = j).$$

En conditionnant par la durée du service et en utilisant les propriétés du processus de Poisson, on obtient

$$a_j = \int_0^{+\infty} \frac{(\lambda s)^j}{j!} e^{-\lambda s} f(s) ds.$$

On notera $G_A(z) = \mathbb{E} z^Y$ la fonction génératrice associée aux Y_n.

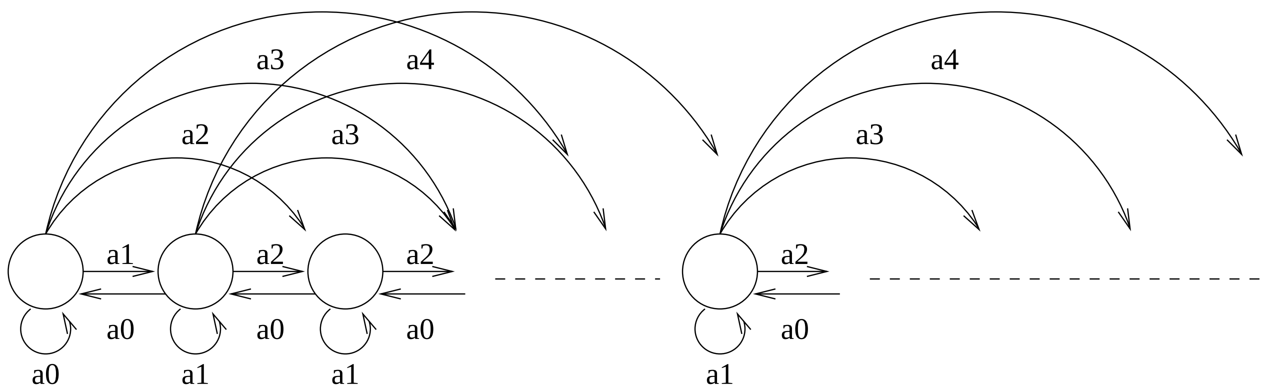

Figure 1: Graphe de la chaîne de Markov incluse $\{X_n\}_{n\in\mathbb{N}}$

La matrice de transition de $\{X_n\}$ est donc

$$\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_0 & a_1 & a_2 & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & a_0 & a_1 & a_2 & \cdots & \cdots & \cdots \\ \vdots & \ddots & \ddots & \ddots & & & & \\ \vdots & & & & & & & \\ \vdots & & & & & & & \end{pmatrix}.$$

Comme $a_i > 0$ car le processus d'arrivée est un processus de Poisson, la chaîne de Markov est apériodique et irréductible. Soit $\pi = (\pi_0, \cdots, \pi_n, \cdots)$ le vecteur de probabilité stationnaire. Les équations de Chapman-Kolmogorov s'écrivent :

$$\pi_k = \pi_0 a_k + \sum_{j=1}^{k+1} \pi_j a_{k+1-j} \quad k \geq 0.$$

Pour résoudre ces équations, on passe par la série génératrice $G(z) = \sum \pi_k z^k$. En multipliant l'équation ci-dessus par z^k et en faisant la somme de tous les équations on obtient :

$$G(z) = \pi_0 G_A(z) + \frac{1}{z}(G(z) - \pi_0) G_A(z).$$

Ceci se réécrit

$$G(z) = \pi_0 \frac{G_A(z)}{1 - \frac{G_A(z)-1}{z-1}}.$$

Le calcul de π_0 est obtenu en utilisant le fait que $G(1) = 1$. Donc en faisant tendre z vers 1 dans l'expression précédente et en utilisant le fait que $G'_A(1) = \lambda \mathbb{E}S$ on tire

$$G(z) = (1-\rho) \frac{G_A(z)}{1 - \frac{G_A(z)-1}{z-1}}.$$

En utilisant l'expression de la série génératrice par la transformée de Laplace du temps de service

$$G_A(z) = \mathcal{L}_S(\lambda(1-z)),$$

on obtient

Théorème 1 (Formule de Pollaczek-Khintchine) *La serie génératrice du nombre de clients à l'état stationnaire, aux instants de départ des clients d'une file d'attente $M/GI/1$ est donnée par la formule :*

$$G(z) = (1-\rho) \frac{\mathcal{L}_S(\lambda(1-z))}{1 - \frac{\mathcal{L}_S(\lambda(1-z))-1}{z-1}}.$$

La difficulté est de passer de la distribution stationnaire de X_n à la distribution stationnaire de X_t. Une propriété importante de la file d'attente $M/GI/1$ est que les clients qui arrivent voient le système à l'état stationnaire. Cette propriété, notée $PASTA$ (Poisson Arrival see Time Average), peut être démontrée dans un cadre beaucoup plus général [**?**] page 54, [**?**] page 293. Comme conséquence de cette propriété, G est la fonction génératrice de X_t à l'état stationnaire.

Cette formulation de la fonction génératrice permet de calculer le nombre moyen de clients dans la file à l'état stationnaire. En effectuant un développement limité en $z = 1$ on obtient :

Corollaire 1
Le nombre moyen de clients dans une file $M/GI/1$ à l'état stationnaire est donné par

$$\mathbb{E}N = \rho + \frac{\rho^2(1 + C_S^2)}{2(1-\rho)},$$

avec $\rho = \lambda \mathbb{E}S$ et C_S^2 le coefficient de variation du temps de service

$$C_S^2 = \frac{Var(S)}{(\mathbb{E}S)^2}.$$

On calcule alors pour différentes distributions en utilisant la valeur de la variance du temps de service.

File $M/D/1$ Dans ce cas la variance du temps de service est nulle, $C_S^2 = 0$ et

$$\mathbb{E}N = \frac{\rho}{1-\rho}(1 - \frac{\rho}{2}).$$

File $M/M/1$ Dans le cas d'une loi exponentielle, $C_S^2 = 1$ et

$$\mathbb{E}N = \frac{\rho}{1-\rho}.$$

On retrouve les résultats obtenus par l'approche markovienne.

File $M/E_k/1$ La loi du temps de service est une loi d'Erlang de paramètre k, de temps moyen de service 1 et de $C_S^2 = \frac{1}{k}$.

$$\mathbb{E}N = \frac{\rho}{1-\rho}\left(1 - \frac{\rho}{2}(1 - \frac{1}{k})\right).$$

La figure suivante montre les écarts lorsque l'on fait varier le coefficient de varaition de 0 à 1 pour des lois d'Erlang de moyenne 1 (c'est à dire des lois $\gamma(k,k)$).

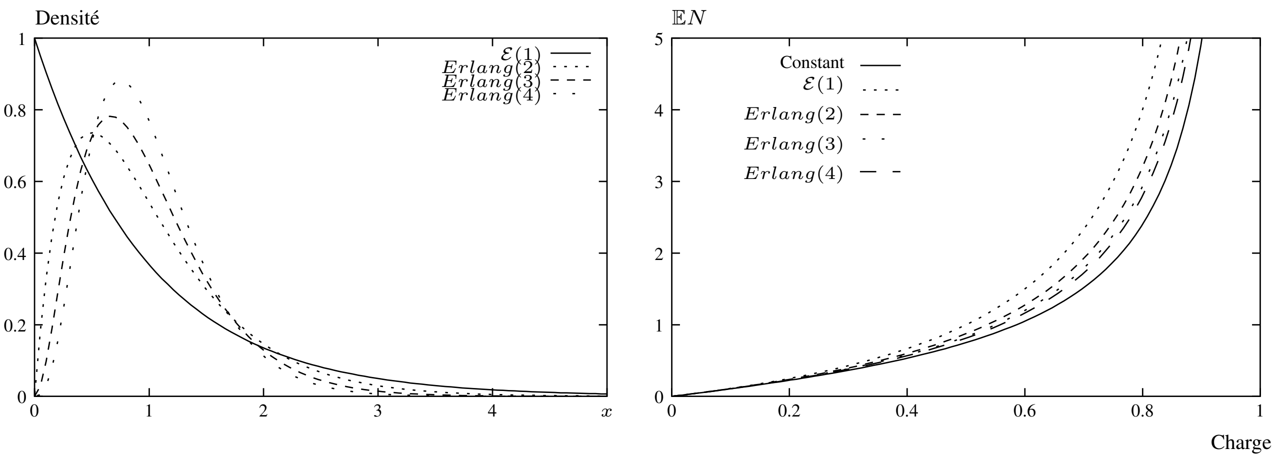

Figure 2: Densité des temps de service et courbes du nombre moyen de clients fonction de la charge.

Il faut noter que le nombre moyen de clients est minimum pour la file à temps de service constant (voir le chapitre sur la comparaison de performances). De plus, lorsque le coefficient de variation est inférieur à 1 le modèle markovien donne une borne supérieure sur le nombre moyen de clients. On se place donc dans un cas "pessimiste" pour l'évaluation de performances. De plus, ces courbes donnent une indication sur l'erreur faite entre un modèle markovien et un modèle déterministe. Cette erreur n'est souvent pas significative par rapport aux erreurs de modélisation et d'estimation des paramètres.