

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
UNIVERSITE AKLI MOAND OULHADJE-BOUIRA



Faculté des Sciences et des Sciences Appliquées
Département de Mathématiques

Mémoire de fin d'étude

Présenté par :

Drider Ibrahim

Rahim Ismahane

En vue de l'obtention du diplôme de **Master** en :

Filière : Mathématiques

Option : Recherche Opérationnelle

Thème :

**Inégalités stochastique pour le réseau
{M/G/1/1 → /G/1/1} avec rappels constants**

Devant le jury composé de :

BOUDREF Mohamed Ahmed
ALEM Lala Maghnia
TAKHEDMIT Baya
HOCINE Safia

MCB
MAA
MCB
MAA

Univ. de Bouira
Univ. de Bouira
Univ. de Béjaia
Univ. de Béjaia

Président
Promotrice
Examinatrice
Examinatrice

Année Universitaire 2018/2019

Table des matières

Introduction générale	5
1 Systèmes et réseaux de files d'attente	7
1.1 systèmes de files d'attente	7
1.1.1 Description du modèle d'attente classique	7
1.1.2 Analyse mathématique d'un système de files d'attente	7
1.1.3 Types de modèles	8
1.1.4 Classification des files d'attente	9
1.1.5 Discipline de service	9
1.1.6 Notation des modèles de files d'attente	9
1.1.7 Mesures de performance de file d'attente	10
1.2 Modèles de files d'attente	11
1.2.1 Le Modèles M/M/1	11
1.2.2 Le Modèles M/G/1	12
1.3 Systèmes de files d'attente avec rappels	13
1.3.1 Description d'un système de files d'attente avec rappels	14
1.4 Modèle d'attente M/G/1 avec rappels	15
1.4.1 Description du modèle	15
1.4.2 Chaîne de Markov induite	15
1.4.3 Distribution stationnaire de la chaîne de Markov induite	16
1.4.4 Distribution stationnaire de l'état du système	17
1.4.5 Décomposition stochastique pour le système M/G/1 avec rappels	20
1.4.6 Mesures de performance :	20
1.5 Réseaux de files d'attente	21
1.5.1 Les caractéristiques d'un réseau de files d'attente	21
1.5.2 Réseau mono ou multi-classe	21
1.5.3 Les réseaux mixtes	23
1.5.4 Les réseaux de Jackson	23
1.5.5 Les réseaux BCMP	24
1.5.6 Quelques exemples	26
conclusion	27
2 les ordres stochastiques dans les réseaux de file d'attente	28
2.1 Propriétés générales des ordres partiels	28
2.1.1 Ordre stochastique	29
2.1.2 Ordre convexe	30
2.1.3 Ordre en transformée de Laplace	32
2.1.4 Ordre en fonctions génératrices	33
2.1.5 Relations entre les ordres partiels	33
2.2 Modèles stochastiques et monotonie	34

2.2.1	Modèles stochastiques	34
2.2.2	Propriétés de monotonie	34
2.3	Comparabilité et monotonie des processus markoviens homogènes	35
2.3.1	Opérateurs monotones et comparables	35
2.3.2	Conditions de monotonie et de comparabilité	37
2.4	Distributions non-paramétriques	37
	Conclusion	38
3	Inégalités stochastique pour le réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants	
	39	
3.1	Description mathématique du modèle	39
3.2	La chaîne de Markov induite	40
3.3	Notations	40
3.4	Monotonie de la chaîne de Markov incluse	41
3.5	Comparaison des opérateurs de transition	44
3.6	Inégalités stochastiques des distributions stationnaires du nombre de clients dans le système	48
	Conclusion	50
	Conclusion générale	51
	Bibliographie	53

Remerciements

Nous remercions plus précisément notre promotrice, Madame Alem Lala Maghnia pour son aide et le temps qu'elle nous a consacré durant le déroulement du mémoire.

Nous exprimons toute notre gratitude à monsieur Boudref Mohamed Ahmed, pour avoir accepté de présider le jury de ce mémoire.

Nos remerciements chaleureux s'adressent également à Takhedmit Baya et Hocine Safia pour d'avoir accepté d'examiner ce travail.

Enfin, nous ne pouvons oublier de remercier tous les membres de nos famille pour leurs soutiens et leurs encouragements, particulièrement nos parents.

Dédicaces

Ibrahim :

Je dédie ce travail à :

- Mes chers parents : je vous souhaite une longue vie,
- Mes chères sœurs : je vous aime tous,
- Toute ma famille,
- Mes amis.

Ismahane :

Je dédie ce mémoire à :

- Mes chers parents : que Dieu vous protège et vous garde pour moi,
- Ma chère grand-mère, source de sagesse : je te souhaite une longue vie,
- Mon cher frère : je te souhaite une bonne chance dans ta vie,
- Mes chères soeurs : je vous aime toutes, une belle vie pour vous,
- Mes amis et tous mes collègues.

Introduction générale

La théorie des files d'attente tire son origine des recherches de l'ingénieur danois Agner Krarup Erlang entre 1909 et 1920. Ce dernier étudiait le concept de la file d'attente des systèmes téléphoniques dans les centres d'appels de Copenhague. Depuis, plusieurs mathématiciens se sont intéressés aux files d'attente et ont développé les modèles mathématiques de cette théorie.

Plusieurs modèles de files d'attente classiques ont été étudiés depuis Erlang, et plusieurs formules "élégantes" ont été élaborées et proposées comme étant des solutions analytiques de certains types de problèmes. Cependant, cette théorie classique s'est très vite montrée inefficace face à des systèmes réels de plus en plus complexes. Par conséquent, les phénomènes de répétitions de demandes du service ont poussé certains chercheurs à étendre le modèle d'attente classique à celui dit avec rappels.

La théorie analytique des modèles d'attente avec rappels s'avère d'une portée limitée en raison de la complexité des résultats connus. En effet, dans la majorité des cas, on se retrouve confronté à des systèmes d'équations dont la résolution est complexe ou possédant des solutions qui ne sont pas facilement interprétables afin que le praticien puisse en bénéficier. Par ailleurs, on peut citer le degré de difficulté pour l'obtention de certaines caractéristiques dans quelques modèles tels que les modèles de files d'attente avec rappels et vacances, avec rappels et priorité, avec rappels constants et blocages, avec rappels de distribution générale ayant deux types de clients.....etc, Cette difficulté réside essentiellement dans l'utilisation des inverses des transformées de Laplace-Stieltjes et des distributions marginales.

Pour pallier à toutes ces difficultés, les chercheurs ont recouru aux méthodes d'approximation qui permettent d'avoir des estimations quantitatives et/ou qualitatives pour certaines mesures de performance. C'est pour toutes ces raisons, qu'on s'intéresse, dans notre étude, d'une manière particulière à la méthode de comparaison stochastique.

L'idée générale de cette méthode est de borner un système complexe par un nouveau système, plus simple à résoudre et fournissant des bornes qualitatives pour ces mesures de performance.

Les propriétés de monotonie et de comparabilité sont bien documentées dans la littérature. L'une de ces propriétés a été étudiée par [14]. Les auteurs ont utilisé la théorie générale des ordres stochastiques pour l'étude des propriétés de monotonie du système M/G/1 avec rappels suivant des ordres stochastiques donnés. En particulier, ils ont déterminé des bornes simples pour le nombre moyen de clients servis durant une période d'activité. Boualem et al. [1] ont étudié quelques problèmes de comparabilité et de monotonie pour l'analyse du système M/G/1 avec rappels constants et vacances du serveur en utilisant la théorie générale des ordres stochastiques. Oukid et Aïssani [20] ont obtenu des bornes stochastiques pour les périodes d'activité et d'inactivité du serveur. Récemment, Alem et all. [16, 17] ont établi quelques approximations qualitatives pour le système

le système $M_1/M_2/G_1/G_2/1$ avec rappels à communication bidirectionnelle et le système $M/G/1$ avec rappels avec rappels a communication bidirectionnelle ayant "n" types d'appels sortant..., en utilisant les propriétés qualitatives des chaînes de Markov par rapport aux taux d'arrivée, aux distributions du temps de service et aux paramètres de rappels. Le lecteur peut également consulter les références [3, 19] traitant la même problématique.

Le but de notre travail est d'appliquer les méthodes de comparaison stochastiques, pour étudier les propriétés de monotonie du réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants .

Ce mémoire est structuré de la manière suivante :

Dans le premier chapitre, on aborde en premier lieu généralités sur les systèmes de files d'attente d'une manière générale et le système $M/G/1$ avec rappels en particulier. Et dans la deuxième partie on cite quelques types de réseaux de files d'attente.

Dans le deuxième Chapitre, on donne un aperçu sur la notion des ordres partiels usuels (ordre stochastique, convexe et de Laplace), ainsi que des éléments sur la théorie de comparabilité des processus stochastiques.

Le troisième Chapitre, est consacré à l'étude de la monotonie de la chaîne de Markov incluse du réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants ainsi on détermine les bornes stochastiques pour la distribution stationnaires de modèle considéré en s'inspirant des travaux de Alem et al. (2019) [16, 17], et on termine par une conclusion générale et une bibliographie.

Chapitre 1

Systemes et reseaux de files d'attente

1.1 systemes de files d'attente

On parle de phénomène d'attente chaque fois que certaines unités appelées "clients" se présentent d'une manière aléatoire à des "stations" afin de recevoir un service dont la durée est généralement aléatoire. La théorie des files d'attente est un formalisme mathématique qui permet de mener des analyses quantitatives à partir de la donnée des caractéristiques du flux d'arrivées et des temps de service.

1.1.1 Description du modèle d'attente classique

Une file d'attente ou queue est un système stochastique composé d'un certain nombre (fini ou non) de places d'attente d'un ou plusieurs serveurs et bien sûr de clients qui arrivent, attendent, se font servir selon des règles de priorité données et quittent le système. La description précédente d'une file d'attente, dont une représentation schématique est donnée en figure (1.1) ne saurait capturer toutes les caractéristiques des différents modèles que comptent la littérature, mais elle identifie les éléments principaux permettant la classification de la grande majorité des files d'attente simples



FIGURE 1.1 – Un système de files d'attente

1.1.2 Analyse mathématique d'un système de files d'attente

L'étude mathématique d'un système de files d'attente se fait généralement par l'introduction d'un processus stochastique, défini de façon appropriée. On s'intéresse principalement au nombre de clients $N(t)$ se trouvant dans le système à l'instant $t (t \geq 0)$.

En fonction des quantités qui définissent le système, on cherche à déterminer :

- Les probabilités d'état $P_n(t) = P(N(t) = n)$, qui définissent le régime transitoire du processus stochastique $\{N(t), t \geq 0\}$. Il est évident que les fonctions $P_n(t)$ dépendent de l'état initial ou de la distribution initiale du processus.

- Le régime stationnaire du processus stochastique qui est défini par :

$$\pi_n = \lim_{t \rightarrow \infty} P_n(t) = \mathbb{P}(N(\infty) = n) = \mathbb{P}(N = n), (n = 0, 1, 2, \dots)$$

$\{\pi_n\}_{n \geq 0}$ est appelée distribution stationnaire du processus $\{N(t), t \geq 0\}$. Le calcul explicite du régime transitoire s'avère généralement pénible, voire impossible, pour la plupart des modèles donnés, On se contente donc de déterminer le régime stationnaire.

1.1.3 Types de modèles

Modèles markoviens

Les modèles markoviens de files d'attente sont des systèmes où les deux quantités stochastiques principales "les temps des inter-arrivées" et "la durée de service" sont des variables aléatoires indépendantes, exponentiellement distribuées. La propriété "sans mémoire" de la loi exponentielle facilite l'analyse de ces modèles.

Processus de Poisson : Un processus de Poisson est un processus stochastique markovien à espace d'état discret et à temps continu, tel que les temps d'inter-arrivées sont des v.a. indépendantes et identiquement distribuées selon une loi exponentielle.

Processus de naissance et de mort : Un processus de naissance et de mort est une généralisation du processus de Poisson. Dans ce processus, les seules transitions possibles à partir d'un état e_n , sont vers les états voisins e_{n-1} et e_{n+1} .

L'étude mathématique d'un système de files d'attente (S.F.A.) se fait souvent par l'introduction d'un processus stochastique défini de façon appropriée de telle sorte à ce qu'il caractérise l'état du système. Le processus stochastique souvent considéré est le processus $\{X(t)\}_{t \geq 0}$: "nombre de clients se trouvant dans le système à l'instant t".

Modèles non markoviens

En l'absence de l'exponentialité ou lorsque l'on s'écarte de l'hypothèse d'exponentialité de l'une des deux quantités stochastiques : le temps des inter-arrivées et la durée de service, ou en prenant en compte certaines spécificités des problèmes par introduction de paramètres supplémentaires, on aboutit à un modèle non markovien. La combinaison de tous ces facteurs rend l'étude mathématique du modèle très délicate, voire impossible. on essaye alors de se ramener à un processus de Markov judicieusement choisi à l'aide de l'une des méthodes d'analyse suivantes :

1. **Méthode des étapes d'Erlang** : Son principe est d'approximer toute loi de probabilité ayant une transformation de Laplace rationnelle par une loi de Cox (mélange de lois exponentielles), cette dernière possède la propriété d'absence de mémoire par étape.
2. **Méthode de la chaîne de Markov induite** :Élaborée par Kendall, et souvent utilisée, elle consiste à choisir une suite d'instantanés $1, 2, 3, \dots, n$ (déterministes ou aléatoires) tels que la chaîne induite $\{N_n, n \geq 0\}$, où $N_n = N(n)$, soit markovienne et homogène.
3. **Méthode des variables supplémentaires** : Elle consiste à compléter l'information sur le processus $\{N(t), t \geq 0\}$ de telle manière à lui donner le caractère markovien. Ainsi, on se ramène à l'étude du processus $\{N(t), A(t_1), A(t_2), \dots, A(t_n), t \geq 0\}$. Les variables $A(t_k), k \in \{1, 2, \dots, n\}$ sont dites supplémentaires.
4. **Méthode des événements fictifs** :Le principe est d'introduire des événements fictifs qui permettent de donner une interprétation probabiliste aux transformées de Laplace et aux variables aléatoires décrivant le système étudié
5. **simulation** : C'est un procédé d'imitation artificielle d'un processus réel effectué sur ordinateur. Elle nous permet d'étudier les systèmes les plus complexes, de prévoir leurs comportements et de calculer leurs caractéristiques. Les résultats obtenus ne sont qu'approximatifs,

mais peuvent être utilisés avec une bonne précision. Cette technique se base sur la génération de variables aléatoires suivant les lois gouvernant le système.

1.1.4 Classification des files d'attente

La classification des files d'attente simples se base principalement sur trois éléments :

- le processus stochastique décrivant la durée de temps séparant deux arrivées consécutives des clients dans le système,
- le nombre de serveurs,
- la loi probabiliste décrivant la durée du temps de service des clients,

A ces trois éléments, il faut parfois ajouter d'autres caractéristiques comme le nombre de places d'attente, le nombre total de clients existants ou encore, les règles spécifiant l'ordre de traitement des clients.

1.1.5 Discipline de service

La discipline de service détermine l'ordre dans lequel les clients sont rangés dans la file et y sont retirés pour recevoir un service. Les disciplines les plus courantes sont :

- FIFO (first in, first out) ou PAPS (premier arrivé, premier servi),
- LIFO (last in, first out) ou (dernier arrivé, premier servi),
- RANDOM (aléatoire) : Le prochain client qui sera servi est choisi aléatoirement dans la file d'attente,
- Priorité relative : Un client accède au service selon sa priorité. La file est gérée par ordre de priorité de la plus forte à la plus faible,
- Priorité absolue : Le service d'un client est interrompu lorsqu'un client de priorité supérieure se présente devant la file d'attente. Le client dont ce service est interrompu est remis en tête de la file.

1.1.6 Notation des modèles de files d'attente

Dans la théorie des file d'attente, la notation de Kendall (premièrement proposée par D.G.Kendall en 1953) est un système standard pour décrire les caractéristiques essentielles d'un modèle de files d'attente. La notation de Kendall a la forme $A/B/s/N/M/D$ où

- A : nature du processus des arrivées,
- B : nature du processus de service,
- s : nombre de serveurs,
- N : capacité d'accueil de la file d'attente,
- M : taille de la population,
- D : discipline de la file.

Dans la description des processus d'arrivée et de service, les symboles les plus courants sont :

- M : loi Exponentielle (memoryless),
- E : loi d'Erlang,
- Γ : loi Gamma,
- D : loi Déterministe (temps d'inter-arrivées ou de service constant),
- G : loi Générale (quelconque).

La forme abrégé : $A/B/s$ signifie que N et M sont infinies.

1.1.7 Mesures de performance de file d'attente

L'analyse d'un modèle de file d'attente a pour objectif de déterminer quantitativement et qualitativement le fonctionnement du système en question. Pour cela, il faut définir les critères et les mesures afin d'atteindre cet objectif [13].

Utilisation ρ : Si le système de file d'attente est constitué d'un seul serveur, alors

$$\rho = \frac{\text{taux d'arrivée}}{\text{taux de service}}$$

L'utilisation d'une file d'attente avec plusieurs serveurs (m) est la fraction moyenne des serveurs actifs. Puisque $m\mu$ est le taux de service global, on a :

$$\rho = \frac{\lambda}{m\mu}$$

Avec :

λ : est le nombre moyen d'arrivées de clients dans le système par unité de temps.

μ : est le nombre moyen de clients servis par unité de temps.

m : est le nombre de serveurs.

On dit que la file d'attente est stable si et seulement si $\rho < 1$ [11].

Temps de réponse T : Également appelé temps de séjour, correspond au temps total passé dans le système de file d'attente. Est la différence entre sa date de départ et sa date d'arrivée [11].

Le temps de réponse T_n d'un client :

$$T_n = d_n - a_n$$

Avec : d_n l'instant de départ et a_n est l'instant d'arrivée [5].

Temps d'attente W : Correspond au temps passé dans une file d'attente en attendant d'être servi, Temps de service : Correspond au temps durant lequel, le client est servi.

$$\text{Temps de réponse} = \text{temps d'attente} + \text{temps de service}$$

Longueur de file d'attente Q : c'est le nombre de clients en attente dans la file.

Nombre de clients dans le système N : c'est le nombre de clients présent dans le système de file d'attente [11]

On a :

$$N = \sum_{k=1}^{\infty} k\pi_k, k=1,2,\dots$$

π_k : est la probabilité d'existence de k clients dans le système de file d'attente.

$$N = \lambda T$$

et de même pour

$$Q = \lambda W$$

1.2 Modèles de files d'attente

1.2.1 Le Modèles M/M/1

La file M/M/1 est la file d'attente markovienne la plus simple, on n'a qu'un seul serveur, où le processus d'arrivée suit un processus de Poisson avec le paramètre $\lambda > 0$, et le temps de service pour chaque client qui suit une loi exponentielle de paramètre $\mu > 0$

a. Régime transitoire

Le processus $\{X(t)\}_{t \geq 0}$: "nombre de clients dans le système à l'instant t ", est markovien. Les équations différentielles de Kolmogorov de ce processus sont de la forme :

$$\begin{cases} p'_0(t) = -\lambda p_0(t) + \mu p_1(t) \\ p'_n(t) = -(\lambda + \mu)p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t), n \in \mathbb{N} \end{cases} \quad (1.1)$$

Ce système d'équations permet de calculer les probabilités d'états $p_n(t)$ en faisant appel aux équations de Bessel et si l'on connaît les conditions initiales (i. e. $X(0)$).

b. Régime stationnaire

Lorsque t tend vers l'infini dans le système (3.6), il est aisé de vérifier $\lim_{t \rightarrow \infty} p_n(t) = \pi_n$ existent et que $\lim_{t \rightarrow \infty} p'_n(t) = 0$, $\forall n \in \mathbb{N}$ Ainsi, à la place d'un système d'équations différentielles, on obtient un système d'équations linéaires et homogènes :

$$\begin{cases} \mu \pi_1 & = \lambda \pi_0 \\ \lambda \pi_{n-1} + \mu \pi_{n+1} & = (\lambda + \mu) \pi_n, n \in \mathbb{N} \end{cases} \quad (1.2)$$

En ajoutant la condition de normalisation " $\sum_{n=0}^{\infty} \pi_n = 1$ " et sous la condition d'ergodicité du système " $\rho = \frac{\lambda}{\mu} < 1$ ", nous obtenons les probabilités stationnaires :

$$\pi_n = \pi_0 \left(\frac{\lambda}{\mu}\right)^n = \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^n, \forall n \in \mathbb{N}$$

Ainsi, le régime stationnaire du système M/M/1 est gouverné par une loi géométrique

c. Mesures de performance de la file M/M/1

► Le nombre moyen de clients dans le système :

$$\begin{aligned} L &= \sum_{n=0}^{\infty} n \pi_n = (1 - \rho) \sum_{n=0}^{\infty} n \rho^n \\ L &= \frac{\rho}{1 - \rho} \\ L &= \frac{\lambda}{\mu - \lambda}. \end{aligned}$$

► Le nombre moyen de clients dans la file :

$$\begin{aligned}
 L_q &= \sum_{n=0}^{\infty} (k-1)\pi_k = \sum_{n=0}^{\infty} (k-1)(1-\rho)\rho^k \\
 L_q &= (1-\rho)\rho^2 \sum_{n=0}^{\infty} (k-1)\rho^{k-2} \\
 L_q &= (1-\rho)\rho^2 \frac{1}{(1-\rho)^2} \\
 L_q &= \frac{\rho^2}{1-\rho}.
 \end{aligned}$$

► Nombre moyen de clients en train d'être servis :

$$L_s = 1 - \pi_0 = \rho.$$

► Temps moyen qu'un client passe dans le système :

$$\begin{aligned}
 W &= \frac{L}{\lambda} = \frac{\rho}{\lambda(1-\rho)} \\
 W &= \frac{1}{\mu - \lambda}.
 \end{aligned}$$

► Temps moyen de service :

$$W_s = \frac{1}{\mu}.$$

► Temps moyen d'attente :

$$W_q = W - W_s = \frac{\lambda}{\mu(\mu - \lambda)}.$$

1.2.2 Le Modèles M/G/1

Les clients arrivent dans le système selon un processus de Poisson de taux $\lambda > 0$. De ce fait, le temps entre deux arrivées successives suit une loi exponentielle de moyenne $\frac{1}{\lambda}$. Le service est assuré par un seul serveur. A l'arrivée d'un client, si le serveur est libre, le client sera pris en charge immédiatement. Dans le cas contraire, il rejoint la file d'attente (de capacité illimitée et discipline FIFO) les durées de service Y sont des variables aléatoires indépendantes et identiquement distribuées de loi générale dont la fonction de répartition et $B(x)$ la transformée de Laplace-Stieltjes $\bar{f}(z)$ et l'espérance mathématique $\mathbb{E}[Y] = \frac{1}{\mu}$.

a. Chaîne de Markov induite

Nous introduisons le processus stochastique $N(t), t \geq 0$ qui n'est pas un processus de Markov. Pour le rendre markovien, nous utiliserons la méthode des chaînes de Markov induites. Soit le processus à temps discret $N_n = N(\xi_n), n \geq 1$ où ξ_n est l'instant où le $n^{\text{ème}}$ client a fini son service et quitte le système. Vérifions que cette suite de variables définit bien une chaîne de Markov.

Soient les (A_n) des variables aléatoires indépendantes et identiquement distribuées telles que A_n est le nombre de clients arrivants pendant le $n^{\text{ième}}$ service avec la distribution

$$\mathbb{P}(A_n = i) = a_i = \int_0^{\infty} \left(\frac{\lambda t}{i!}\right)^i \exp(-\lambda t) dB(t),$$

Déterminons l'équation fondamentale de la chaîne :

$$N_{n+1} = \begin{cases} N_n - 1 + A_{n+1}, & \text{si } N_n \geq 1 \\ A_{n+1}, & \text{si } N_n = 0 \end{cases}, n \geq 1$$

Soit la variable aléatoire

$$\delta = \begin{cases} 1 & \text{si } N_n > 0 \\ 0 & \text{si } N_n = 0 \end{cases}$$

alors l'équation fondamentale de la chaîne devient :

$$N_{n+1} = N_n - \delta_n + A_{n+1}. \quad (1.3)$$

Il est évident que N_{n+1} dépend de N_n et A_{n+1} seulement et non pas de N_{n-1}, N_{n-2}, \dots . D'où la suite $\{N_n, n \geq 1\}$ est une chaîne de Markov induite du processus $\{N(t), t \geq 0\}$ avec les probabilités de transitions $\mathbb{P}(N_{n+1} = j / N_n = i) = p_{ij}$ qui s'expriment de la manière suivante :

$$\begin{cases} p_{0j} = a_j, & \text{si } j \geq 0 \\ p_{ij} = a_{j-i+1}, & \text{si } 0 \leq i \leq j+1 \\ p_{ij} = 0, & \text{ailleurs.} \end{cases}$$

b. Mesures de performance

► Nombre moyen de clients dans le système

$$L = \mathbb{E}(X) = \rho + \frac{\rho^2 + \lambda^2 + \text{Var}(Y)}{2(1 - \rho)}.$$

► Le nombre moyen de client dans la file d'attente

$$L_q = L - \rho = \frac{\rho^2 + \lambda^2 + \text{Var}(Y)}{2(1 - \rho)}.$$

► Temps moyen de séjour d'un client dans le système

$$W = \frac{L}{\lambda} = \frac{1}{\mu} + \left(\frac{\lambda \left(\text{Var}(Y) + \frac{1}{\mu^2} \right)}{2(1 - \rho)} \right).$$

► Temps moyen d'attente d'un client

$$W_q = \frac{L_q}{\lambda} = \frac{\lambda \left(\text{Var}(Y) + \frac{1}{\mu^2} \right)}{2(1 - \rho)}.$$

1.3 Systèmes de files d'attente avec rappels

Dans la théorie des files d'attente classique, il est supposé qu'un client qui ne peut pas obtenir son service immédiatement dès son arrivée, rejoint la file d'attente ou quitte le système définitivement. Les systèmes de files d'attente développés tentent de prendre en considération des phénomènes de répétition de demandes de service, et ceci après une durée du temps aléatoire. Un tel système est connu comme «système de files d'attente avec rappels».

Pour identifier un système de files d'attente avec rappels, on a besoin des spécifications suivantes : la nature stochastique du processus des arrivées, la distribution du temps de service, le nombre de serveurs qui composent l'espace de service, la capacité et discipline d'attente ainsi que la spécification concernant le processus de répétition d'appels.

1.3.1 Description d'un système de files d'attente avec rappels

Modèle général

Le modèle décrit ci-dessus est entièrement général.

Pour désigner les systèmes avec rappels, on utilise la notation de Yang et Templeton [22]

$A/B/s/m/O/H$, où A et B décrivent respectivement la distribution du temps inter-arrivées et la distribution du temps de service, s représente le nombre de serveurs dans le système, m est le nombre de positions d'attente plus le nombre de serveurs, et O est la capacité de l'orbite, (elle peut-être supprimée lorsque sa capacité est infinie). La séquence $H = \{H_i, i \geq 0\}$ est la fonction de persévérance, où H_i est la probabilité qu'un client fasse une $(i + 1)^{\text{ème}}$ tentative de rappel, après une $i^{\text{ème}}$ tentative échouée. Si tous les clients sont persévérants, $H_j = 1$, pour tout j , le symbole H pourra être également supprimé.

Maintenant, considérons un système de file d'attente, de flux des arrivées primaires poissonniennes de paramètre λ , et ayant s ($s \geq 1$) serveurs identiques et indépendants, de distribution de service générale, de fonction de répartition $B(x)$. Il y a dans ce système $m - s$ ($m \geq s$) positions d'attente. À l'arrivée d'un client, s'il y a un ou plusieurs serveurs libres, le client sera immédiatement pris en charge. Sinon, s'il y a une position libre, le client rejoint la file d'attente. Dans le cas contraire, si le client trouve tous les serveurs et les positions d'attente occupées, il quitte le système définitivement avec une probabilité $1 - H_0$, ou quitte le système temporairement avec une probabilité H_0 et rappelle ultérieurement. La capacité O de l'orbite peut-être finie ou infinie. Dans le cas où O est finie, si l'orbite est pleine, le client quitte le système définitivement. Chaque client de l'orbite forme un processus d'arrivées secondaires de taux μ et il est traité de la même manière qu'un client primaire. S'il trouve un serveur ou une position libre, il reçoit immédiatement le service. Par contre, si tous les serveurs et les positions d'attente sont occupés, il quitte le système définitivement avec une probabilité $1 - H_k$ (s'il s'agit de la $k^{\text{ème}}$ tentative infructueuse), ou rejoint l'orbite (si elle n'est pas pleine) avec une probabilité H_k . Le schéma général d'un système avec rappel est donné par la figure suivante :

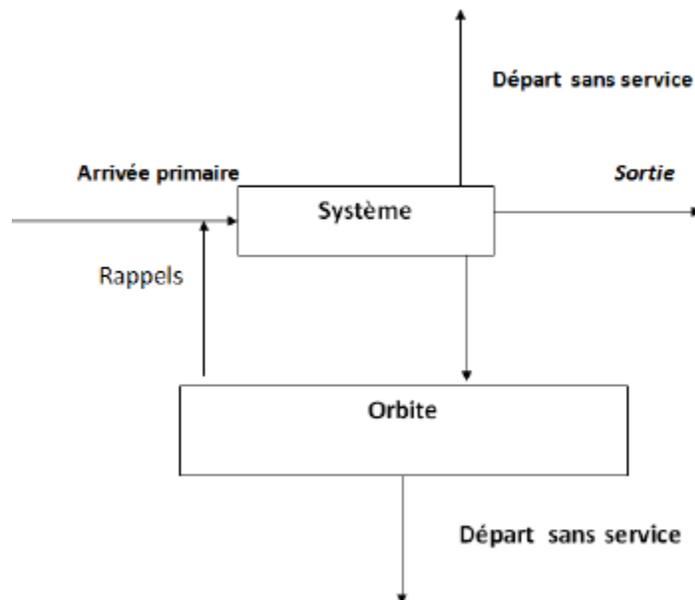


FIGURE 1.2 – Schéma général d'un système d'attente avec rappels.

1.4 Modèle d'attente M/G/1 avec rappels

Le modèle M/G/1 avec rappels est le modèle le plus étudié par les spécialistes. Il existe une littérature abondante sur ses propriétés.

1.4.1 Description du modèle

Considérons un système à un seul serveur, les arrivées des clients dans le système forment un processus de Poisson de taux λ , si un client primaire trouve le serveur disponible, il est immédiatement pris en charge. Sinon, c'est-à-dire, s'il trouve le serveur occupé, il rejoint l'orbite et rappelle ultérieurement jusqu'à ce qu'il trouve le serveur de nouveau disponible. Le temps de service est de loi générale de moyenne $\frac{1}{\beta_1}$, de distribution $B(x)$ et de transformée de Laplace-Stieltjes $\bar{f}(z)$. La durée entre deux rappels successifs d'une même source secondaire est exponentielle de taux μ . La description du système est la suivante :

On suppose que le $(i - 1)^{\text{ème}}$ appel termine son service à l'instant η_{i-1} (les appels sont numérotés dans l'ordre de service) et le serveur devient libre. Même s'il y a des clients dans le système, ils ne peuvent occuper le service immédiatement. Donc le $i^{\text{ème}}$ appel suivant, n'entre en service qu'après un intervalle de temps R_i durant lequel le canal est libre, bien qu'en général il y a des clients qui attendent. A l'instant $\xi_i = \eta_{i-1} + R_i$, le $i^{\text{ème}}$ client débute son service durant un temps S_i .

Tous les rappels qui arrivent durant ce temps de service n'influent pas sur le processus. Alors à l'instant $\eta_i = \xi_i + S_i$, le $i^{\text{ème}}$ client achève son service et le canal devient encore libre et ainsi de suite.

1.4.2 Chaîne de Markov induite

L'état du système à l'instant t peut être décrit par le processus

$$X(t) = (C(t), N_0(t), \varsigma(t), t \geq 0)$$

où :

$C(t)$: la variable aléatoire indiquant l'état du serveur à l'instant t .

$$C(t) = \begin{cases} 0 & \text{si le serveur est libre,} \\ 1 & \text{si le serveur est occupé.} \end{cases}$$

$N_0(t)$: le nombre de clients en orbite à l'instant t .

Si $C(t) = 1$, $\varsigma(t)$ représente le temps de service écoulé du client en service. Soit (q_n) la chaîne de Markov induite aux instants de départ, où $q_n = N_0(\varsigma_n)$ représente le nombre de clients en orbite après le $n^{\text{ème}}$ départ, dont l'équation fondamentale est :

$$q_{n+1} = q_n - \sigma_{q_n} + \nu_{n+1}$$

où :

ς_{n+1} est le nombre de clients qui arrivent dans le système durant le service du $(n + 1)^{\text{ème}}$ client. La distribution de ν_{n+1} est donnée par :

$$a_i = \mathbb{P}(\nu = i) = \int_0^{\infty} \frac{(\lambda x)^i}{i!} \exp(-\lambda x) dB(x) \quad \text{où } a_i > 0, i \geq 0$$

Si $\lim_{n \rightarrow \infty} \nu_n = \nu$ et $\mathbb{E}(\nu) = \tau$ alors

$$A(z) = \sum_{i \geq 0} a_i z^i = \int_0^{\infty} \exp^{-x(\lambda - \lambda z)} dB(x) = \tilde{B}(\lambda - \lambda z)$$

σ_n : est la variable de Bernoulli définie par :

$$\sigma_{q_n} = \begin{cases} 1 & \text{si le } (n+1)^{\text{ème}} \text{ client servi provient de l'orbite,} \\ 0 & \text{sinon} \end{cases}$$

Elle dépend de q_n et sa distribution est donnée par :

$$\mathbb{P}(\sigma_{q_n} = 1/q_n = i) = \frac{i\mu}{\lambda + i\mu}, \quad \mathbb{P}(\sigma_{q_n} = 0/q_n = i) = \frac{\lambda}{\lambda + i\mu}$$

Les probabilités de transition à une étape de l'état i vers l'état j notées r_{ij} sont données par :
 $r_{ij} = \mathbb{P}(q_{n+1} = j/q_n = i) = a_{j-i} \frac{\lambda}{\lambda + i\mu} + a_{j-i+1} \frac{i\mu}{\lambda + i\mu} \quad (\forall j \geq 0 \text{ et } \forall 0 \leq i \leq j)$

En effet

$$\begin{aligned} r_{ij} &= \mathbb{P}(q_{n+1} = j/q_n = i) = \mathbb{P}(q_n - \sigma_{q_n} + v_{n+1} = j/q_n = i) \\ &= \mathbb{P}(v_{n+1} = j - i + \sigma_{q_n}/q_n = i) \\ &= \mathbb{P}(v_{n+1} = j - i/q_n = i, \sigma_{q_n} = 0) \mathbb{P}(\sigma_{q_n} = 0/q_n = i) + \mathbb{P}(v_{n+1} = j - i + 1/q_n = i, \sigma_{q_n} = 0) \mathbb{P}(\sigma_{q_n} = 1/q_n = i) \\ &= \mathbb{P}(v_{n+1} = j - i) \mathbb{P}(\sigma_{q_n} = 0/q_n = i) + \mathbb{P}(v_{n+1} = j - i + 1) \mathbb{P}(\sigma_{q_n} = 1/q_n = i) \\ &= a_{j-i} \left[\frac{\lambda}{\lambda + i\mu} \right] + a_{j-i+1} \left[\frac{i\mu}{\lambda + i\mu} \right]. \end{aligned}$$

Condition d'existence du régime stationnaire

L'accroissement moyen de la chaîne est :

$$\begin{aligned} \mathbb{E}[q_{n+1} - q_n/q_n = i] &= \mathbb{E}[\nu_{n+1} - \sigma_{q_n}/q_n = i] \\ &= \mathbb{E}[v_{n+1}] - \mathbb{E}[\sigma_{q_n} = 1/q_n = i] \\ &= \rho - \frac{i\mu}{\lambda + i\mu}. \end{aligned}$$

Si $\rho < 1$, alors $\lim_{i \rightarrow \infty} \mathbb{E}[q_{n+1} - q_n/q_n = i] = \rho - 1 < 0 \implies$ la chaîne est donc érgodique.

Si $\rho \geq 1$, alors $\lim_{i \rightarrow \infty} \mathbb{E}[q_{n+1} - q_n/q_n = i] = \rho - \frac{i\mu}{\lambda + i\mu} \geq 1 - \frac{i\mu}{\lambda + i\mu} = \frac{\lambda}{\lambda + i\mu} > 0$. Puisque la chaîne est bornée inférieurement par la chaîne induite du système M/G/1 classique, donc la chaîne n'est pas érgodique (elle est transitoire).

1.4.3 Distribution stationnaire de la chaîne de Markov induite

Soit $\rho < 1$ et $\pi = (\pi_0, \pi_1, \dots)$ la distribution stationnaire de la chaîne de Markov induite. Il ne sera généralement pas possible de trouver la distribution stationnaire π , mais nous pouvons calculer la fonction génératrice correspondante $\pi(z)$.

Théoreme 1.4.1. *La distribution stationnaire de la chaîne de Markov induite possède la fonction génératrice suivante :*

$$\pi(z) = \frac{(1 - \rho)(1 - z)A(z)}{A(z) - z} \exp \left[\frac{\lambda}{\mu} \int_1^x \frac{1 - A(u)}{A(u) - u} du \right]$$

où $A(z) = \bar{f}(\lambda - \lambda z)$.

1.4.4 Distribution stationnaire de l'état du système

Supposons que $\rho < 1$. La méthode des variables supplémentaires, permet de trouver la distribution stationnaire du système à partir de l'état du serveur et la taille de l'orbite.

Elle consiste à décrire le processus des arrivées comme processus de Markov avec dépendance de l'état de paramètre λ_n quand $\{C(t), N_0(t)\}$ est dans l'état (i, n) et à appliquer les schémas récursifs. L'état du système peut être décrit par le processus suivant :

$$X(t) = \begin{cases} N_0(t) & \text{si } C(t) = 0 \\ \{C(t), N_0(t), \xi(t)\} & \text{si } C(t) = 1 \end{cases}$$

où $\xi(t)$ est une variable aléatoire supplémentaire à valeurs dans \mathbb{R}_+ , désignant la durée de service écoulé à l'instant t . Notons par

$$p_{0n} = \lim_{t \rightarrow \infty} P(C(t) = 0, N_0(t) = n)$$

$$p_{1n}(x) = \lim_{t \rightarrow \infty} \frac{d}{dx} P(C(t) = 1, \xi(t) \leq x, N_0(t) = n)$$

les probabilités p_{0n} et $p_{1n}(x)$ vérifient le système d'équations de balance :

$$(\lambda + n\mu)(p_{0n}) = \int_0^{\infty} p_{1n}(x)f(x)dx, n \geq 0$$

$$p'_{1n}(x) = -(\lambda + f(x)p_{1n}(x) + \lambda p_{1n-1}(x)), n \geq 1$$

$$p_{1n}(0) = \lambda p_{0n} + (n+1)\mu p_{0n+1}, n \geq 0$$

ou $b(x) = \frac{B'(x)}{1-B(x)}$ avec $B(x) \neq 1$ est l'intensité instantanée du service étant donné que la durée écoulée est égale à x .

Soient les fonctions génératrices :

$$P_0(z) = \sum_{n=0}^{\infty} z^n p_{0n} \text{ et } P_1(z, x) = \sum_{n=0}^{\infty} z^n p_{1n}(x)$$

Le système d'équations de balance devient

$$\begin{cases} \lambda \sum_{n=0}^{\infty} z^n p_{0n} + \mu \sum_{n=0}^{\infty} z^n n p_{0n} = \int_0^{\infty} \sum_{n=0}^{\infty} z^n p_{1n}(x) f(x) dx \\ \sum_{n=0}^{\infty} z^n p'_{1n}(x) = -(\lambda + f(x)) \sum_{n=0}^{\infty} z^n p_{1n}(x) + \lambda \sum_{n=0}^{\infty} z^n p_{1n-1}(x) \\ \sum_{n=0}^{\infty} z^n p_{1n}(0) = \lambda \sum_{n=0}^{\infty} z^n p_{0n} + \mu \sum_{n=0}^{\infty} z^n (n+1) p_{0n+1} \end{cases}$$

D'où

$$\lambda P_0(z) + \mu z P'_0(z) = \int_0^{\infty} P_1(z, x) b(x) dx \quad (1.4)$$

$$P'_1(z, x) = (\lambda z - \lambda - b(x)) P_1(z, x) \quad (1.5)$$

$$P_1(z, 0) = \lambda P_0(z) + \mu P'_0(z) \quad (1.6)$$

De (1.5), on a

$$P_1(z, x) = P_1(z, 0)[1 - B(x)] \exp(-(\lambda - \lambda z)x)$$

Donc (1.4), devient

$$\begin{aligned}
\lambda P_0(z) + \mu P_0'(z) &= \int_0^{\infty} P_1(z, 0)[1 - B(x)] \exp(-(\lambda - \lambda z)x) b(x) dx \\
&= P_1(z, 0) \int_0^{\infty} b'(x) \exp(-(\lambda - \lambda z)x) dx \\
&= P_1(z, 0) \int_0^{\infty} \exp(-(\lambda - \lambda z)x) dB(x) \\
&= P_1(z, 0) \tilde{B}(\lambda - \lambda z) \\
&= P_1(z, 0) A(z).
\end{aligned} \tag{1.7}$$

A partir des équation (1.7) et (1.6), on a

$$\begin{aligned}
P_1(z, 0) A(z) &= \lambda P_0(z) + \mu z \left(\frac{P_1(z, 0)}{\mu} - \frac{\lambda}{\mu} P_0(z) \right) \\
P_1(z, 0) (A(z) - z) &= (\lambda - \lambda z) P_0(z) \\
P_1(z, 0) &= \frac{(\lambda - \lambda z)}{(A(z) - z)} P_0(z) \\
\Rightarrow P_1(z, x) &= \frac{\lambda - \lambda z}{A(z) - z} P_0(z) [1 - B(x)] \exp(-(\lambda - \lambda z)x)
\end{aligned}$$

En intégrant cette équation, et en utilisant la formule suivante

$$\int_0^{\infty} \exp(-sx) [1 - B(x)] dx = \frac{1 - \tilde{B}(s)}{s}$$

On obtient

$$\begin{aligned}
P_1(z) &= \int_0^{\infty} P_1(z, x) dx = \frac{(\lambda - \lambda z)}{(A(z) - z)} P_0(z) \int_0^{\infty} [1 - B(x)] \exp(-(\lambda - \lambda z)x) dx \\
&= \frac{(\lambda - \lambda z)}{(A(z) - z)} P_0(z) \frac{1 - \tilde{B}(\lambda - \lambda z)}{\lambda - \lambda z} \quad \text{avec } \tilde{B}(\lambda - \lambda z) = A(z) \\
\Rightarrow P_1(z) &= P_0(z) \frac{1 - A(z)}{A(z) - z}
\end{aligned}$$

De (1.7) et (1.4), on peut obtenir

$$\lambda P_0(z) + \mu z P_0'(z) = A(z) [\lambda P_0(z) + \mu P_0'(z)] \tag{1.8}$$

$$\mu [A(z) - z] P_0'(z) = \lambda [1 - A(z)] P_0(z) \tag{1.9}$$

Considérons La fonction analytique $f(z) = A(z) - z$ est positive, croissante et pour $z \in [0, 1]$ $\rho < 1 : z < A(z) < 1$.

Théoreme 1.4.2. si $\rho = \lambda\beta_1 < 1$, le système est en régime stationnaire et les fonctions génératrices de la distribution conjointe de l'état du serveur et de la taille de l'orbite sont données par :

$$P_0(z) = \sum_{n=0}^{\infty} z^n p_{0n} = (1 - \rho) \exp \left[\frac{\lambda}{\mu} \int_1^z \frac{1 - A(u)}{A(u) - u} du \right]$$

$$P_1(z) = \sum_{n=0}^{\infty} z^n p_{1n} = \frac{1 - A(z)}{A(z) - z} P_0(z).$$

L'équation 1.9 devient

$$P'_0 = \frac{\lambda}{\mu} \left[\frac{1 - A(z)}{A(z) - z} \right] P_0(z)$$

La résolution de cette équation nous donne

$$P_0(z) = \exp \left[\frac{\lambda}{\mu} \int_1^z \frac{1 - A(u)}{A(u) - u} du \right]$$

$$P_1(z) = \left[\frac{1 - A(z)}{A(z) - z} \right] P_0(z)$$

De plus, $P_1(1) = \frac{\rho}{1-\rho} P_0(1)$. Et vu que, $P_0(1) + P_1(1) = 1$ on obtient $P_1(1) = \rho$ et $P_0(1) = 1 - \rho$. Par conséquent, la distribution marginale du nombre de serveurs occupés s'exprime de la manière suivante :

$$P_0 = \lim_{t \rightarrow \infty} P(C(t) = 0) = (1 - \rho) \quad \text{et} \quad P_1 = \lim_{t \rightarrow \infty} P(C(t) = 1) = \rho$$

La fonction génératrice de la distribution marginale de la taille de l'orbite est définie par :

$$\begin{aligned} P(z) = P_0(z) + P_1(z) &= P_0(z) + \frac{1 - A(z)}{A(z) - z} P_0(z) \\ &= \left[\frac{1 - A(z)}{A(z) - z} \right] P_0(z) \\ &= \frac{1 - z}{A(z) - z} P_0(z) \\ &= \frac{(1 - z)(1 - \rho)}{A(z) - z} \exp \left[\frac{\lambda}{\mu} \int_1^{\infty} \frac{1 - A(u)}{A(u) - u} du \right]. \end{aligned}$$

Et la fonction génératrice de la distribution de l'état stationnaire du nombre de clients dans le système est :

$$\begin{aligned} Q(z) = P_0(z) + zP_1(z) &= P_0(z) + z \left[\frac{1 - A(z)}{A(z) - z} \right] P_0(z) \\ &= \left[1 + \frac{z - zA(z)}{A(z) - z} \right] P_0(z) \\ &= \frac{A(z)(1 - z)}{A(z) - z} P_0(z) \\ &= \frac{(1 - z)(1 - \rho)A(z)}{A(z) - z} \exp \left[\int_1^{\infty} \frac{1 - A(u)}{A(u) - u} du \right] \end{aligned}$$

Qui est la fonction génératrice de la distribution de la chaîne induite.

1.4.5 Décomposition stochastique pour le système M/G/1 avec rappels

Parmi les approches permettant d'étudier les systèmes de files d'attente avec rappels, on rencontre celles basées sur la propriété de décomposition stochastique que peut posséder un modèle. Le concept général de la propriété de décomposition stochastique d'un système d'attente M/G/1 avec rappels est défini de la manière suivante : le nombre de clients se trouvant dans le système à une date aléatoire est distribué comme la somme de deux variables aléatoires indépendantes ou plus, l'une de ces variables représente le nombre de clients se trouvant dans le système M/G/1 classique à une date aléatoire (le serveur est toujours disponible).

La propriété de décomposition stochastique présente diverses applications pratiques dans le modèle M/G/1 avec rappels parmi eux :

- Estimation de la vitesse de convergence vers le système M/G/1 classique. μ tend vers ∞

$$D = \sum_{i=0}^1 \sum_{n=0}^{\infty} |p_{in}(\mu) - p_{in}(\infty)| = o\left(\frac{1}{\mu}\right)$$

- Estimation de la distance entre les distributions stationnaires du modèle M/G/1 avec rappels et le modèle M/G/1 classique correspondant : $2(1 - \rho)(1 - N_0(\mu)) < D < 2(1 - N_0(\mu))$ où

$$N_0(\mu) = \exp \left\{ -\frac{\lambda}{\mu} \int_0^1 \frac{1 - \bar{f}(\lambda - \lambda x)}{f(\lambda - \lambda x) - x} dx \right\}$$

$N_0(\mu)$: le nombre de client en orbite à la date t .

1.4.6 Mesures de performance :

Les caractéristiques du système M/G/1 avec rappels sont données dans l'article de Yang et Templeton [22]. Dans ce qui suit on note

β_i le $i^{\text{ème}}$ moment du temps de service. Avec

$$\rho = \lambda\beta_1.$$

:

► Nombre moyen de clients dans le système

$$L = Q'(1) = \rho + \frac{\lambda^2 \beta_2}{2(1-\rho)} + \frac{\lambda\rho}{\mu(1-\rho)}.$$

► Nombre moyen de clients en l'orbite

$$L_0 = L - \rho = \frac{N_0}{\lambda} = \frac{\lambda^2 \beta_2}{2(1-\rho)} + \frac{\lambda\rho}{\mu(1-\rho)}.$$

► Temps moyen d'attente d'un client

$$W = \frac{L_0}{\lambda} = \frac{\lambda\beta_2}{2(1-\rho)} + \frac{\rho}{\mu(1-\rho)}.$$

► Nombre moyen de rappels par client (d'après la formule de Little)

$$R = W\mu = \frac{\lambda\mu\beta_2}{2(1-\rho)} + \frac{\rho}{1-\rho}.$$

1.5 Réseaux de files d'attente

La modélisation d'un système à l'aide d'une seule file d'attente n'offre qu'un champ d'application restreint. Bien souvent, un client a besoin de recevoir plusieurs traitements consécutifs et de différentes natures avant de quitter un système.

Les réseaux de files d'attentes permettant de modéliser de telles situations et correspondent à des systèmes composés de plusieurs files d'attentes reliées entre elles. Lorsqu'un client quitte une file, il peut se diriger vers une nouvelle station du réseau ou sortir définitivement du système. Ces décisions sont dictées par des règles de routage qui peuvent être déterministes ou stochastiques.

Un réseau de file d'attente est simplement un système composé d'une ou plusieurs files d'attentes reliées entre elles. Les clients (dans les cas les plus simples, tous identiques) une fois leur service termine dans une station (file), se déplace vers une autre station ou quittent le système selon des règles de routage.

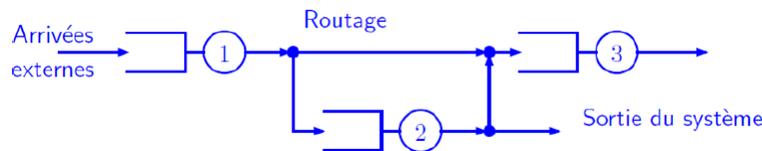


FIGURE 1.3 – Exemple d'un réseau de files d'attente.

1.5.1 Les caractéristiques d'un réseau de files d'attente

Un réseau de files d'attente est caractérisé par [21] :

- Le chemin suivi par les clients dans le système : ce chemin peut être statique (chaque client va de serveur en serveur suivant une route prédéterminé), soit dynamique : après le service dans une station, le client a le choix entre plusieurs autres stations pour poursuivre son traitement. Ce choix est éventuellement pondéré par des fonctions de probabilité,
- Les classes de clients : plusieurs classes de clients peuvent être présentes dans le système. Tous les clients appartenant à une même classe suivront le même chemin (décrit statiquement ou dynamiquement). Par contre, à l'intérieur de la station, les différentes classes restent en compétition pour recevoir leur traitement.

1.5.2 Réseau mono ou multi-classe

Un réseau parcouru par une seule classe de clients est appelé réseau mono-classe. Contrairement au réseau multi-classe dans lequel circulent plusieurs classes. Chaque classe de clients peut être soit ouverte soit fermée.

a.1 Réseau mono-classe ouvert

On appelle réseau mono-classe ouvert un réseau avec une seule classe de client où les clients arrivent de l'extérieur du système, séjournent pour recevoir un ou plusieurs services, puis quittent définitivement le système. Par conséquent, le nombre de clients est infini. La figure 1.4 montre un exemple de réseau mono-classe ouvert.

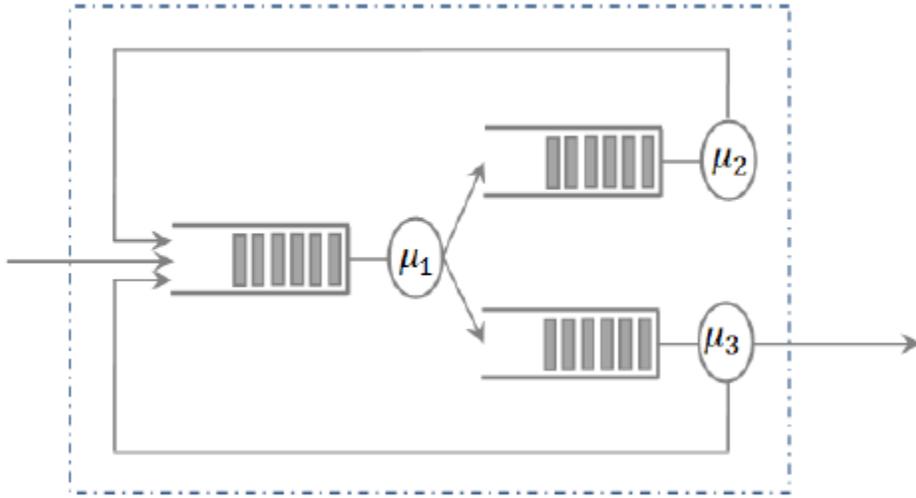


FIGURE 1.4 – Réseau mono-classe ouvert

a.2 Réseau mono-classe fermés

Un réseau est dit mono-classe fermé lorsque tous les clients appartiennent à la même classe et leur nombre en entrée du système est constant. Il n'y a ni de départ ni d'arrivée de clients. La figure 1.5 montre un exemple de réseau mono-classe fermé.

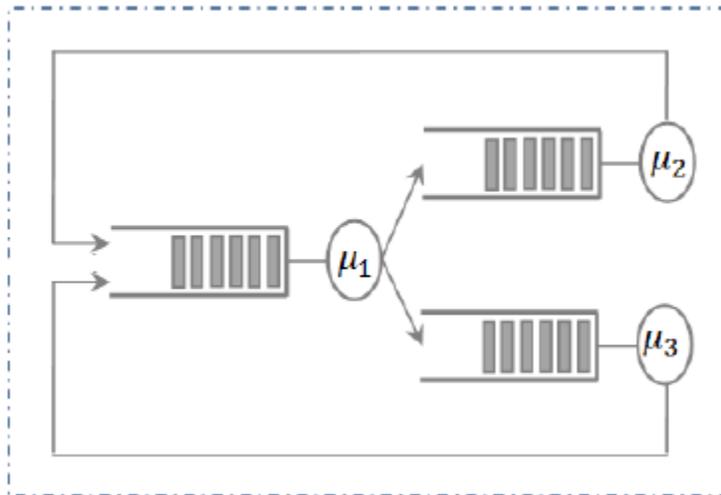


FIGURE 1.5 – Réseau mono-classe fermés

b. Réseaux multi-classes

Les réseaux de files d'attente peuvent être parcourus par différentes classes de clients. Soit r le nombre de classes de clients. Ces différentes classes se distingueront par :

- Des processus d'arrivée différents (si le réseau est ouvert),
- Des comportements différents à chaque station (service et discipline de service),
- Des routages différents dans le réseau.

On est alors amené à caractériser chaque classe r :

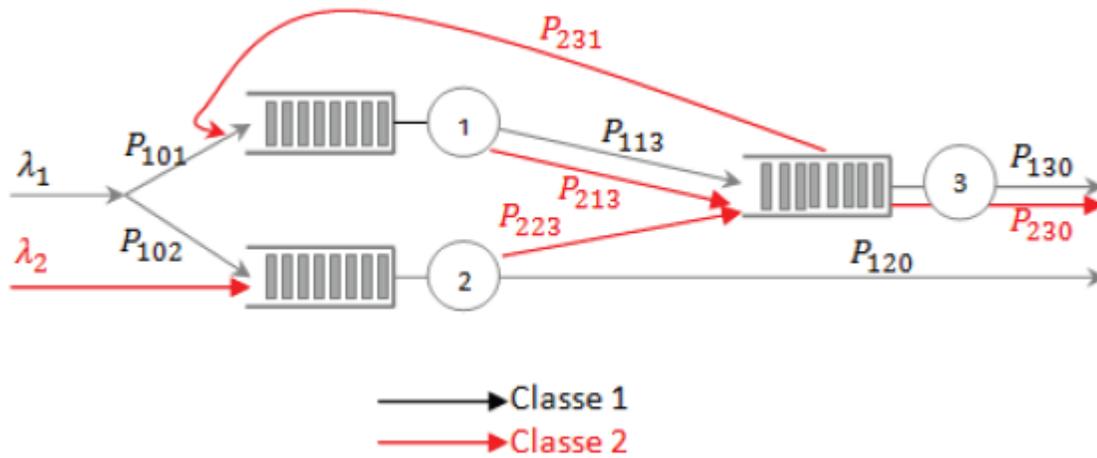


FIGURE 1.6 – Réseau multi-classe

- Pour un réseau ouvert, les clients de classe r arrivent de l’extérieur dans la file i suivant un processus de Poisson de paramètre $\lambda_i r$.
- Pour un réseau fermé, N est le nombre total de clients de classe r ,
- Le routage des clients. Si on se limite au routage probabiliste, on définit P_{rij} la probabilité pour qu’un client de classe r qui quitte la station i , se rende à la station j . (Si i ou j est égale à 0, cela fait référence à « l’extérieur » d’un réseau ouvert) [18, 12].

1.5.3 Les réseaux mixtes

Ce sont des réseaux ouverts pour certaines classes et fermés pour d’autres classes.

1.5.4 Les réseaux de Jackson

Par réseaux de Jackson, on entend un ensemble de files d’attente reliées entre elles de façon quelconque, la distribution des temps de service étant exponentiellement distribués de taux μ_i et le processus des arrivées de l’extérieur étant poissonnien de taux λ_i . Les files sont de capacité illimitée de telle sorte qu’il n’y ait pas de blocage : la discipline de service est FIFO. Les clients (appartenant tous à la même classe) sont acheminés dans le réseau suivant des probabilités de routage p_{ij} : probabilité en sortant de la file i d’aller vers la file j . Si le réseau est ouvert possède une source notée 0 qui permet à des clients venant de l’extérieur d’entrer dans le réseau par la file j avec la probabilité p_{0j} . De plus, il existe une station puit dénotée $K+1$ par laquelle les clients quittent le système. Le flux total des clients arrivant dans le système forme un processus de Poisson de paramètre λ .

$$\sum_{j=1}^{k+1} p_{ij} = 1, \quad i = 0, \dots, k.$$

Si le réseau est fermé, les stations 0 et $K+1$ sont confondues : il n’y a pas de clients entrant dans le système ni de clients pouvant en partir. Le nombre de clients reste constant dans un réseau fermé. Soit n_1, n_2, \dots, n_K le nombre de clients respectivement dans la file 1, 2, ..., K . Le comportement du réseau est totalement défini par les valeurs de $P(n_1, n_2, \dots, n_K, t)$: la probabilité d’être au temps t dans l’état n_1, n_2, \dots, n_K , que nous appellerons probabilité jointe, et la probabilité marginale $P(n_i, t)$ d’avoir n_i clients dans la file i au temps t .

Un réseau de file d'attente sera dit en équilibre s'il existe un état stationnaire. Dans ce cas, nous ôterons le temps t des probabilités jointes et marginales.

Théoreme 1.5.1. *Un réseaux de Jackson ouvert, stable et formé de n files possède une distribution stationnaire unique donnée par*

$$P(n_1, n_2, \dots, n_k) = \prod_{i=1}^k P(n_i). \quad (1.10)$$

*Ainsi, en régime stationnaire, un réseau de Jackson se comporte comme autant de files de type $M/M/m$ isolées recevant leurs clients selon des processus de Poisson indépendants les uns des autres. Une distribution de la forme (1.10) est dite à forme produit et les réseaux associés, **des réseaux à forme produit.***

1.5.5 Les réseaux BCMP

Les réseaux BCMP (Basket, Chandy, Muntzet, Palacios) sont une extension des réseaux de Jackson. Ils gardent la solution à forme produit à l'état d'équilibre, mais ils autorisent plusieurs classes de clients et des disciplines de service différente de FIFO. Les réseaux considérés par BCMP doivent accomplir les caractéristiques suivantes [9].

A. Discipline des files d'attente

Les caractéristiques des réseaux BCMP mènent à définir les quatre types de files à forme produit suivantes [10].

- Type-1 :-/M/m - FCFS,
- Type-2 : -/G/1 - PS,
- Type-3 : -/G/ ∞ (IS),
- Type-4 : -/G/1 - LCFS PR.

Dans notre cas d'étude nous allons utiliser le premier type car notre réseau est un ensemble de file d'attente type M/M/1- FCFS.

B. Distribution des temps de service

Les temps de service des noeuds type FCFS doivent être exponentiellement distribués et indépendants des différentes classes (c'est-à-dire : $\mu_{i1} = \mu_{i2} = \dots = \mu_{iR} = \mu_i$).

C. Processus des arrivées

Le processus d'arrivée suit le processus de poisson, lorsque tous les clients arrivent au réseau d'une seule source avec le taux d'arrivée λ . où λ dépend du nombre de client dans le réseau. Les clients sont distribués dans les noeuds du réseau selon les probabilités de routage.

Nous allons utiliser les notations suivantes pour définir ce type de réseau :

- R : nombre des classes de clients dans le réseau,
- K_{ir} : nombre des clients dans le noeud «i» de la classe «r»,
- μ_{ir} : taux de service du noeud «i» de la classe «r»,
- $R_{ir,js}$: la probabilité que le client du noeud «i» de la classe «r» est transféré au noeud «j» de la classe «s» (probabilité de routage),
- $R_{0,js}$: la probabilité dans un réseau ouvert qu'un client de l'extérieur d'un réseau entre dans le noeud «j» de la classe «s»,
- $R_{ir,0}$: la probabilité dans un réseau ouvert qu'un client de la classe «r» quitte le réseau après avoir été servi dans le noeud «i»,
- $\lambda_{0,ir}$: taux d'arrivée de l'extérieur du système au noeud «i» de la classe «r».
- λ_{ir} : taux d'arrivée des client au noeud «i» de la classe «r»,
- C_q : chaîne des classes dans le cas ou les clients changent de classe (Dans notre réseau C_q est le nombre de classe de clients car les clients ne changent pas de classes dans le réseau).

Théoreme 1.5.2. (BCMP). [9] Pour un réseau BCMP ouvert, le calcul de sa probabilité stationnaire est :

$$\pi(k_1, \dots, k_n) = \prod_{i=1}^n \pi_i(k_i) \text{ avec } \pi_i(k_i) = (1 - \rho_i) \rho_i^{k_i}$$

Il faut s'assurer que la condition de stabilité : (1.11) est vérifiée pour toutes les files du réseau.

$$\rho_i < 1 \quad (1.11)$$

Pour déterminer les mesures de performances d'un réseau BCMP, il faut suivre ces cinq étapes :

1. Calculer le taux de visite e_{ir} pour toutes les files d'attente $i = 1, \dots, N$ et toutes les classes $r = 1, \dots, R$.
2. Calculer l'utilisation du serveur de chaque file d'attente (ρ_i),
3. Calculer les autres mesures de performances ($\bar{N}_{ir}, \bar{T}_{ir}, \bar{W}_{ir}$ et \bar{Q}_{ir}),
4. Calculer la probabilité stationnaire de chaque file d'attente $\pi(k_i)$,
5. Calculer la probabilité stationnaire pour tout le réseau $\pi(k_1, \dots, k_n)$.

Mesures de performance des réseaux BCMP ouverts

Dans cette section nous allons décrire les formules analytiques [9] nécessaires pour le calcul des mesures de performance des réseaux ouverts multi-classes.

► Taux de visite e_{ir}

Dans un réseau BCMP ouvert multi-classe, le taux de visite ou le nombre moyen de passage à la station i de la classe r est obtenu comme suit :

$$e_{ir} = R_{0,ir} + \sum_{s \in C_q, i=1, \dots, N} e_{is} R_{js,ir}$$

Telque : $r \in C_q, i = 1, \dots, N, r = 1, \dots, R$.

► Utilisation ρ_{ir}

L'utilisation de chaque serveur d'une file d'attente $i = 1, \dots, n$ appartenant à une classe : $r = 1, \dots, R$ est donnée par la formule suivante

$$\rho_{ir} = \lambda_r \frac{e_{ir}}{\mu_{ir}}$$

L'utilisation du serveur $i = 1, \dots, n$ par toutes les classes est calculée par la formule :

$$\rho_i = \sum_{r=1}^R \rho_{ir}$$

► Nombre moyen des clients \bar{N}_{ir}

Le nombre moyen des clients dans le système (serveur+buffer) $i = 1, \dots, n$ de la classe $r = 1, \dots, R$ est donné par la formule suivante :

$$\bar{N}_{ir} = \frac{\rho_{ir}}{1 - \rho_i}$$

► Temps de réponse moyen \bar{T}_{ir}

Le temps de réponse moyen de la file d'attente $i = 1, \dots, n$ de la classe $r = 1, \dots, R$ est obtenu aussi par la formule de Little comme suit :

$$\bar{T}_{ir} = \frac{\bar{N}_{ir}}{\lambda_{ir}}$$

► **Temps d'attente moyen** \bar{W}_{ir}

Si les taux des services sont indépendants alors le temps d'attente moyen est le suivant :

$$\bar{W}_{ir} = \bar{T}_{ir} - \frac{1}{\mu_{ir}}.$$

► **Longueur moyenne de la file d'attente** \bar{Q}_{ir}

La longueur moyenne de la file d'attente (buffer) $i = 1, \dots, n$ de la classe $r = 1, \dots, R$ est calculée par l'équation de Little et elle vaut :

$$\bar{Q}_{ir} = \lambda_{ir} \bar{W}_{ir}$$

1.5.6 Quelques exemples

Réseaux de communication par paquet

Considérons un réseau de communications d'ordinateurs dans lequel on trouve un ensemble d'interfaces IMPs (Interface Message Processors) reliées entre elles par des câbles.

Un ordinateur principal est connecté à l'une de ces interfaces. Si l'ordinateur veut envoyer un message à un autre ordinateur principal, il doit en premier lieu envoyer le message avec l'adresse de destination à l'interface à laquelle il est connecté. L'interface à son tour envoie le message à l'ordinateur destinataire directement si elle y est connectée, ou indirectement via d'autres interfaces. Une interface à laquelle un ordinateur principal est connecté. Les messages arrivent de l'extérieur selon un processus aléatoire. Après la réception du message, l'ordinateur l'envoie immédiatement à son interface. S'il y a un tampon libre, le message est accepté. Dans le cas contraire, le message est rejeté et l'ordinateur doit réessayer une autre fois après une période de temps. S'il existe des tampons libres, le message rejeté sera stocké dans un tampon de l'ordinateur principal. Dans le cas contraire, le message est considéré comme perdu. On peut poser les questions suivantes : Le problème présenté peut être modélisé comme un système avec rappels à serveur unique (interface IMP) possédant des tampons (positions d'attente). Le nombre de tampons de l'ordinateur principal constitue la capacité de l'orbite.

Système mémoire sur les disques magnétiques

Considérons un système mémoire où K unités disques partagent un contrôleur des disques, et transmettent l'information quand elles trouvent ce dernier libre. Les demandes insatisfaites sont répétées après une rotation du disque. Ce système peut être présenté comme un système de files d'attente avec rappels. Le serveur est le contrôleur des disques. L'arrivée dans le système est quasi-aléatoire : le nombre de sources est fini et égal à K . La rotation du disque dans le cas de la répétition de demande peut être considérée comme un intervalle entre deux rappels consécutifs de durée constante.

Problème de réservation

Le problème de réservation est l'exemple le plus simple d'un client qui sollicite une réservation par téléphone dans un restaurant. Il y a une ligne unique qui est consacrée à répondre aux requêtes des réservations. Ainsi, si un client appelle et trouve la ligne occupée, il renouvellera sa tentative après une certaine période de temps aléatoire avec la probabilité p qui, en pratique, est strictement inférieure à 1 car le client ne peut rappeler indéfiniment.

Cet exemple peut être modélisé par une file d'attente M/G/1 avec rappels et avec perte en considérant que le processus d'arrivée des appels est poissonnien. L'étude de ce genre de problèmes permet de prédire le temps d'attente du client, le nombre de clients perdus dû à ce blocage, ..

Conclusion

Chaque file d'attente est caractérisée par son processus d'arrivée, taux de service et la discipline de la file. Dans certains systèmes, un client accomplit son travail en passant par plusieurs serveurs d'où la notion des réseaux de files d'attente.

Les réseaux de files d'attente ont une très grande importance car ils servent à modéliser des systèmes physiques, ils permettent d'évaluer les performances et ils aident à mieux comprendre le comportement de ces systèmes.

Chapitre 2

les ordres stochastiques dans les réseaux de file d'attente

Les ordres stochastiques sont de plus en plus utilisés dans plusieurs domaines de recherche bien établi, qui est toujours en développement intensif et qui offre plusieurs problèmes ouverts. Ils mènent à des méthodes d'approximation robustes et des bornes dans des situations où les modèles stochastiques réels sont trop complexes pour un traitement rigoureux. Les ordres stochastiques sont aussi utiles dans des situations où les distributions fondamentales d'un modèle sont seulement connues partiellement.

Dans ce chapitre nous allons présenter plusieurs théorèmes, résultats, propriétés, corollaires et définitions sur la théorie des ordres stochastiques, de D. Stoyan [4].

2.1 Propriétés générales des ordres partiels

On appelle un ordre partiel, noté " \prec ", une relation binaire définie sur un ensemble D d'éléments a, b, c, \dots , satisfaisant les trois axiomes :

1. $a \prec a$ (réflexivité),
2. si $a \prec b$ et $b \prec c$ alors $a \prec c$ (transitivité),
3. si $a \prec b$ et $b \prec a$ alors $a = b$ (antisymétrie).

Notons que $a \prec b$ est équivalent à dire que $b \succ a$.

Cette section est consacrée à quelques propriétés de l'ordre partiel défini sur l'ensemble D de toutes les fonctions de répartition de variables aléatoires réelles (ou bien l'un de ses sous-ensembles).

Pour les deux variables aléatoires X et Y de fonctions de répartition F et G (respectivement) on a par convention :

$$F \prec G \Leftrightarrow X \prec Y.$$

On suppose que deux variables aléatoires X et Y sont définies sur le même espace de probabilité, alors leurs fonctions de répartition respectives F et G peuvent satisfaire la propriété d'antisymétrie (3) sans pour autant avoir $X = Y$.

Lorsque les variables aléatoires sont dégénérées, certaines propriétés des ordres partiels définies sur D découlent directement des propriétés de l'ordre des nombres réels. Pour cela, on utilisera la distribution de Dirac, notée par $\Theta_c(\cdot)$, définie pour tous les nombres réels comme suit :

$$\Theta_c(x) = \begin{cases} 0, & \text{si } x < c, \\ 1, & \text{si } x \geq c. \end{cases}$$

Définition 2.1.1.

Soit un ordre partiel donné " \prec " défini sur (un sous ensemble de) l'espace \mathcal{D} des fonctions de répartition.

On dit que cet ordre possède la propriété :

- (R) : si $\forall a, b \in \mathbb{R}$ tels que $a \leq b$, alors $\Theta_a < \Theta_b$.
- (E) : si $F \prec G$, alors $m_F \leq m_G$ lorsque les moyennes existent.
- (M) : si $F \prec G$, alors $F^c \prec G^c$, $\forall c > 0$, où $F^c(x) = F(x/c)$, $\forall x$.
- (C) : si $F_1 \prec F_2$ alors $F_1 * G \prec F_2 * G$, où $(F_i * G)(x) = \int_{-\infty}^{\infty} F_i(x-y)dG(y)$, $i = 1, 2$.
- (w) : si F_n et G_n convergent faiblement vers F et G (respectivement) alors :

$$\forall n, F_n \prec G_n \Rightarrow F \prec G.$$

Remarque 2.1.1.

Pour les deux variables aléatoires X et Y :

La propriété (M) assure que :

$$X \prec Y \Leftrightarrow cX \prec cY \text{ pour tout } c \in]0, +\infty[.$$

La propriété (C) assure que :

$$X_1 \prec X_2 \Rightarrow X_1 + Y \prec X_2 + Y,$$

où Y est une variable aléatoire indépendante de X_1 et X_2 . La propriété (E) assure que :

$$X \prec Y \Rightarrow E(X) \geq E(Y).$$

On remarque que la propriété (E) découle des autres propriétés.

Proposition 2.1.1.

Un ordre partiel \prec sur un ensemble (ou bien sur un sous ensemble de) \mathcal{D} qui vérifie les propriétés (R); (M); (C) et (W), vérifie aussi la propriété (E).

Définition 2.1.2.

Pour une classe de fonctions réelles \mathfrak{S}_{\prec} , l'ordre partiel \prec défini sur l'ensemble (ou sur le sous ensemble de) \mathcal{D} est dit généré par \mathfrak{S}_{\prec} si :

$$F \prec G \Leftrightarrow \int_{-\infty}^{+\infty} f(x)dF(x) \geq \int_{-\infty}^{+\infty} f(x)dG(x),$$

pour toute fonction f dans \mathfrak{S}_{\prec} , telle que les intégrales existent.

2.1.1 Ordre stochastique

Définition 2.1.3.

On dit que la variable aléatoire X de fonction de répartition F , est stochastiquement inférieure (ou bien inférieure en distribution) à la variable aléatoire Y de fonction de répartition G , et on note $F \leq_{st} G$, lorsque

$$F(x) \geq G(x), \forall x \in \mathbb{R}.$$

On écrit aussi $X \leq_{st} Y$ (\leq_{st} noté aussi par l'ordre \leq_d).

Dans le cas où X et Y sont des variables aléatoires discrètes prenant des valeurs sur l'ensemble des entiers relatifs \mathbb{Z} , et en notant par $P_i^{(1)} = P\{X = i\}$ et $P_i^{(2)} = P\{Y = i\}$ pour $i \in \mathbb{Z}$, alors

$$X \leq_{st} Y \Leftrightarrow \sum_{j=-\infty}^i P_j^{(1)} \geq \sum_{j=-\infty}^i P_j^{(2)}, i \in \mathbb{Z},$$

ce qui est équivalent à :

$$\sum_{j=i}^{+\infty} P_j^{(1)} \leq \sum_{j=i}^{+\infty} P_j^{(2)}, i \in \mathbb{Z}.$$

Remarquons que l'ordre stochastique \leq_{st} satisfait les axiomes de l'ordre partiel \prec .

Proposition 2.1.2.

Si $F_1 \leq_{st} F_2$, alors il existe deux variables aléatoires X_1 et X_2 définies sur le même espace de probabilité $(\Omega, \mathcal{A}; \mathcal{P})$ pour lesquelles

$$X_1(\omega) \leq X_2(\omega), ; \forall \omega \in \Omega,$$

et

$$P(\omega : X_k(\omega) \leq x) = F_k(x) \text{ pour } k = 1, 2.$$

Notons par $\mathfrak{R}_{st}(\mathbb{R})$ la classe des fonctions réelles non décroissantes, alors la classe $\mathbb{R}_{\leq_{st}}$ des fonctions \leq -monotones est confondue avec la classe $\mathfrak{R}_{st}(\mathbb{R})$, c'est-à-dire $(\mathbb{R})_{\leq_{st}} = \mathfrak{R}_{st}(\mathbb{R})$

Théoreme 2.1.1.

L'inégalité

$$\int_{-\infty}^{+\infty} f(t) dF_1(t) \leq \int_{-\infty}^{+\infty} f(t) dF_2(t), \quad (2.1)$$

est vérifiée pour toute fonction f appartenant à $\mathfrak{R}_{st}(\mathbb{R})$ pour laquelle l'intégrale existe, si et seulement si $F_1 \leq_{st} F_2$. Pour une fonction f donnée, l'inégalité (2.1) est vérifiée pour tout F_1 et F_2 telles que $F_1 \leq_{st} F_2$ uniquement si f est non décroissante.

2.1.2 Ordre convexe

On note par $x_+ = \max(0, x)$.

Définition 2.1.4.

On dit que la variable aléatoire X , de fonction de répartition F , est inférieure en moyenne de vie résiduelle à la variable aléatoire Y , de fonction de répartition G , et on écrit $X \leq_v Y$, ou bien $F \leq_v G$ si et seulement si :

$$\begin{aligned} E((X - x)_+) &= \int_x^{+\infty} (t - x) dF(t) = \int_x^{+\infty} (1 - F(t)) dt \\ &\leq \int_x^{+\infty} (1 - G(t)) dt = E((Y - x)_+), \end{aligned} \quad (2.2)$$

lorsque les espérances (ou bien les intégrales) sont bien définies. Dans le cas discret, on a :

$$X \leq_v Y \Leftrightarrow \sum_{i=k}^{\infty} \sum_{j=i}^{\infty} P_j^{(1)} \leq \sum_{i=k}^{\infty} \sum_{j=i}^{\infty} P_j^{(2)}.$$

Une conséquence immédiate de cette définition :

$$\text{si } F \leq_{st} G \text{ et } E(Y_+) < \infty \text{ alors } F \leq_v G.$$

Théorème 2.1.2.

1. L'inégalité

$$\int_{-\infty}^{+\infty} f(t) dF_1(t) \leq \int_{-\infty}^{+\infty} f(t) dF_2(t), \quad (2.3)$$

est vérifiée pour toute fonction f appartenant à $\mathfrak{R}_v(\mathbb{R})$ pour laquelle les intégrales sont bien définies, si et seulement si $F_1 \leq_v F_2$.

2. Pour une fonction donnée f , l'inégalité (2.3) a lieu pour toutes les fonctions F_1 et F_2 telles que $F_1 \leq_v F_2$ uniquement si f est une fonction convexe et non décroissante.
3. Si $F_1 \leq_v F_2$ et leurs moyennes existent et sont égales, alors l'inégalité (2.3) est vérifiée pour toute fonction convexe f donnée.

Pour deux variables aléatoires X et Y non négatives telles que $X \leq_v Y$ on a

$$E(X^r) \leq E(Y^r), (r \geq 1),$$

lorsque les espérances existent.

En général, pour des variables aléatoires X et Y telles que

$$E(X) = E(Y), \text{ et } X \leq Y,$$

alors,

$$E(X^r) \leq E(Y^r), (r = 2, 4, 6, \dots).$$

Il est intéressant de remarquer que pour deux variables aléatoires telles que X et Y sont non négatives et $X \leq_v Y$, alors l'égalité $E(X^r) = E(Y^r)$ pour tout $r \geq 1$ implique l'égalité $X =_{st} Y$.

En effet

$$E(X^r) = \int_0^{+\infty} r x^{r-1} (1 - F(x)) dx = \int_0^{+\infty} r(r-1) x^{r-2} dx \int_x^{+\infty} (1 - F(y)) dy.$$

Cette propriété est l'analogie de la propriété suivante pour l'ordre stochastique

$$X \leq_{st} Y \text{ et } E(X) = E(Y) \Rightarrow X =_{st} Y.$$

Proposition 2.1.3.

Supposons que les suites de variables aléatoires X_n et Y_n convergent faiblement vers X et Y (respectivement).

Si

$$\begin{aligned} E(X_+) \text{ et } E(Y_+) \text{ sont finis,} \\ E((X_n)_+) \longrightarrow E(X_+) \text{ quand } n \longrightarrow +\infty, \\ E((Y_n)_+) \longrightarrow E(Y_+) \text{ quand } n \longrightarrow +\infty, \end{aligned}$$

et si $X_n \leq_v Y_n$, alors

$$X \leq_v Y$$

2.1.3 Ordre en transformée de Laplace

Transformée de Laplace

Lorsque la variable aléatoire X est du type continu, sa distribution peut être caractérisée par la transformée de Laplace de la densité $f(x)$:

$$\hat{f}(s) = E(e^{-sX}) = \int_0^{\infty} f(x)e^{-sx} dx,$$

où s est une variable complexe. Cette intégrale est définie au moins pour $\text{Re}(s) \geq 0$. La transformée de Laplace est notée aussi $L[f(x)]$.

Propriétés

- Si X et Y sont indépendantes, la transformée de Laplace de $X + Y$ est le produit des transformées de Laplace de X et de Y ,
- $L[f'(x)] = s\hat{f}(s) - f(0)$,
- $L[f''(x)] = s^2\hat{f}(s) - sf(0) - f'(0)$,
- $L[\int_0^x f(u)du] = \frac{\hat{f}(s)}{s}$,
- Si $F(x)$ est la fonction de répartition de X et si $R(x) = 1 - F(x)$ alors

$$\lim_{s \rightarrow 0} \hat{R}(s) = \int_0^{\infty} R(x)dx.$$

Définition 2.1.5.

Pour deux variables aléatoires non négatives X et Y de fonctions de répartition F et G (respectivement), F est dite inférieure par rapport à l'ordre laplacien à G , et on note $F \leq_L G$, si pour tout s positif on a l'inégalité suivante

$$E(\exp(-sX)) = \int_0^{\infty} \exp(-sX)dF(x) \geq \int_0^{\infty} \exp(-sX)dG(x) = E(\exp(-sY)).$$

Il est clair que l'ordre en transformée de Laplace est réflexif, transitif et antisymétrique.

Théoreme 2.1.3.

Soit une fonction f strictement monotone, alors $F \leq_L G$ implique

$$\int_0^{+\infty} f(t)dF(t) \leq \int_0^{+\infty} f(t)dG(t).$$

1. Pour deux variables aléatoires X et Y non négatives, de fonctions de répartition F et G respectivement, telles que $F \leq_L G$ alors, on a l'inégalité suivante :

$$\frac{1 - E(\exp(-sX))}{s} \leq \frac{1 - E(\exp(-sY))}{s},$$

2. Lorsqu'on fait tendre s vers 0, on obtient le résultat suivant :

$$F \leq_L G \Rightarrow E(X) \leq E(Y),$$

lorsque les espérances existent.

Le résultat qui suit donne une caractérisation de l'ordre en transformée de Laplace.

Théoreme 2.1.4.

Soient X et Y deux variables aléatoires quelconques de fonctions de répartition F et G respectivement, alors :

$$F \leq_L G \Leftrightarrow E(f(X)) \leq E(f(Y)),$$

pour toute fonction f strictement monotone, telle que les espérances existent.

2.1.4 Ordre en fonctions génératrices

Soient X et Y deux variables aléatoires non négatives discrètes de fonctions de répartition F et G respectivement. On dit que X est inférieure à Y par rapport à l'ordre en fonctions génératrices, et on note $F \leq_g G$, si et seulement si :

$$E(z^X) \geq E(z^Y),$$

où,

$$E(z^X) = \sum_{n=0}^{+\infty} P(X = n)z^n \text{ et } E(z^Y) = \sum_{n=0}^{+\infty} P(Y = n)z^n, |z| < 1.$$

Cet ordre peut-être déduit de l'ordre laplacien en posant $s = -\ln z$.

2.1.5 Relations entre les ordres partiels

Soient X et Y deux variables aléatoires de fonctions de répartition F et G respectivement. Alors, on a les relations suivantes :

- Si $F \leq_{st} G$ et $E(Y_+) < \infty \Rightarrow F \leq_v G$.
- Si $E(X) = E(\max(0; -x)) < \infty \Rightarrow F \leq_{cv} G$.
- Si $E(X) = E(Y)$, alors $F \leq_{cv} G \Leftrightarrow G \leq_v F$.
- $F \leq_{st} G \Rightarrow F \leq_L G \Rightarrow F \leq_g G$.
- $F \leq_{cv} G \Rightarrow F \leq_L G \Rightarrow F \leq_g G$.
- Si $E(X) = E(Y)$, et $F \leq_v G \Rightarrow G \leq_L F \Rightarrow G \leq_g F$.
- $F \leq_L G \Rightarrow F \leq_g G$.

2.2 Modèles stochastiques et monotonie

2.2.1 Modèles stochastiques

La modélisation est la représentation d'un système réel par un modèle mathématique. On parle de modèle stochastique lorsque les influences aléatoires sont prises en considération. Cette prise en compte du hasard et ses effets permet une meilleure compréhension des phénomènes réels et la résolution efficace de problèmes complexes. Plusieurs disciplines de la recherche opérationnelle reposent sur les probabilités et leurs applications. En particulier, les théories de files d'attente, de fiabilité et de gestion des stocks.

2.2.2 Propriétés de monotonie

Étudier mathématiquement les modèles stochastiques, c'est d'obtenir des estimations des quantités qui, pour un modèle Σ donné, avec une structure spécifique et des distributions F_i des X_i, \dots , décrivent son comportement.

Soit c_Σ une caractéristique dans Σ et soit C_Σ l'ensemble des valeurs possibles de c_Σ .

Pour une structure donnée et une distribution initiale U , c_Σ dépend uniquement des F_i , et on écrit

$$c_\Sigma = c_\Sigma(F_1, F_2, \dots) \in C_\Sigma.$$

Pour des modèles simples, on peut déduire une expression explicite de c_Σ . Cependant, dans plusieurs situations, cela n'est pas possible et les calculs mathématiques peuvent mener à des formules compliquées qui ne peuvent pas être exploitées en pratique.

De telles circonstances nous suggèrent de rechercher les propriétés qualitatives de c_Σ par rapport aux F_i , i.e, la manière avec laquelle c_Σ est affectée par les changements en F_i . Parmi les propriétés qualitatives importantes des modèles stochastiques on trouve la monotonie (i.e, si les F_i croissent dans un certain sens, alors c_Σ croît aussi).

Monotonie interne :

Soit Σ un modèle stochastique constitué de distributions paramétriques $(U, F_1, F_2, \dots) \equiv (U, F)$, où U est la distribution initiale.

On note par C_Σ les indices de performance du système Σ .

Par exemple, pour un système de files d'attente Σ , c_Σ peut-être le temps moyen d'attente virtuel à l'instant t , ou la distribution de probabilité du nombre de clients dans le système à l'instant t , ou bien ses mesures de performance seront calculées à une suite d'instant $(t_n)_{n \in \mathbb{N}^*}$ (déterministe ou aléatoire).

D'une manière plus précise on peut exprimer c_Σ comme suit :

$$C_\Sigma = C_\Sigma(t, U, F_i), \text{ ou bien}$$

$$C_\Sigma(t_n) = C_\Sigma(n) = C_\Sigma(n, U, F_i),$$

On note par \prec l'ordre partiel défini sur c_Σ .

Définition 2.2.1.

L'indice de performance $c_\Sigma(\cdot)$ est non décroissant (resp. non croissant) par rapport à la distribution initiale U si pour tout $t < u$, on a :

$$t < u \Rightarrow c_\Sigma(t) \prec c_\Sigma(u) \text{ (resp. } c_\Sigma(t) \succ c_\Sigma(u)), \quad (2.4)$$

ou bien pour les entiers $m < n$:

$$m < n \Rightarrow c_{\Sigma}(m) \prec c_{\Sigma}(n) \text{ (resp. } c_{\Sigma}(m) \succ c_{\Sigma}(n)),$$

Cette propriété est appelée monotonie interne. D'autres appellations sont utilisées telles que, monotonie temporelle ou intrinsèque. Celles-ci découlent du fait que cette monotonie ne dépend en aucun cas des distributions paramétriques $\{F_i\}$, mais seulement peut-être de la distribution initiale.

Monotonie externe :

On note par \mathcal{D}_k l'ensemble des distributions F_k partiellement ordonnées par l'ordre " $<_k$ " (qui est l'ordre de la $k^{\text{ème}}$ distribution paramétrique), et soit " $<_c$ " l'ordre partiel défini sur c_{Σ} .

Définition 2.2.2.

L'indice de performance c_{Σ} est non décroissant sur \mathcal{D}_k par rapport à l'ordre $<_k$ si pour tout F et G dans \mathcal{D}_k et toute autre distribution paramétrique constante, on a :

$$F <_k G \Rightarrow c_{\Sigma}(F_1, \dots, F_{k-1}, F, F_{k+1}, \dots) <_c (F_1, \dots, F_{k-1}, G, F_{k+1}, \dots),$$

Cette propriété est appelée monotonie externe.

Lorsqu'un système possède la propriété de monotonie externe, les indices de performance des modèles stochastiques, possédant la même structure avec des distributions paramétriques comparables mais différentes, sont comparables.

On peut interpréter la monotonie externe comme suit :

Soient Σ_1 et Σ_2 deux modèles stochastiques ayant la même structure et la même distribution initiale. On dira que ces modèles possèdent la propriété de monotonie externe lorsque pour deux distributions paramétriques F et G dans Σ_1 et Σ_1 respectivement, on a :

$$F \prec G \Rightarrow c_{\Sigma}(F) <_c c_{\Sigma}(G),$$

pour l'indice de performance c_{Σ} .

La monotonie externe est un outil d'une grande importance dans la construction des bornes pour les mesures de performance d'un système donné. Ainsi, la distribution paramétrique F_k peut être bornée par les distributions G_1 et G_2 appartenant à l'ensemble \mathcal{D}_k pour lesquelles :

$$G_1 <_k F_k <_k G_2,$$

alors pour les mesures de performance correspondantes, on obtient :

$$c_{\Sigma}(G_1) <_k c_{\Sigma}(F) <_k c_{\Sigma}(G_2),$$

lorsque les systèmes ont la propriété de la monotonie externe.

2.3 Comparabilité et monotonie des processus markoviens homogènes

2.3.1 Opérateurs monotones et comparables

Soient \mathbb{E}, \mathbb{M} un espace probabilisable et $P_{\mathbb{M}}$ l'ensemble de toutes les mesures de probabilité définies sur M. Soient aussi les opérateurs $\tau, \tau^{(1)}$ et $\tau^{(2)}$ définis de $P_{\mathbb{M}}$ dans $P_{\mathbb{M}}$ et l'ordre partiel " \preceq " défini sur $P_{\mathbb{M}}$.

Définition 2.3.1.

Un opérateur τ est dit \prec -monotone si pour toutes mesures de probabilités $p^{(1)}, p^{(2)}$ appartenant à $P_{\mathbb{M}}$ telles que $p^{(1)} \prec p^{(2)}$, on a

$$\tau p^{(1)} \prec \tau p^{(2)}.$$

L'opérateur $\tau^{(1)}$ est inférieur à $\tau^{(2)}$ si $\tau^{(1)}p \prec \tau^{(2)}p$ pour tout $p \in P_{\mathbb{M}}$ et on écrit

$$\tau^{(1)} \prec \tau^{(2)}.$$

Pour des applications aux processus de Markov homogènes, on s'intéresse à la comparabilité des distributions $p_n^{(1)}$ et $p_n^{(2)}$ définies par

$$p_n^{(k)} = (\tau^{(k)}p^{(k)})_n, \quad k = 1, 2 \text{ et } n \in N^*$$

pour deux distributions initiales $p^{(k)}$ et les opérateurs $\tau^{(k)}$, pour $k = 1, 2$.

Théoreme 2.3.1.

soient $\tau^{(1)}, \tau^{(2)}$ deux opérateurs définis sur $P_{\mathbb{M}}$ et $p^{(1)}, p^{(2)}$ deux mesures de probabilité définies sur \mathcal{M} , alors

$$p^{(1)} \prec p^{(2)} \text{ implique } p_n^{(1)} \prec p_n^{(2)}, \quad \forall n \in N^*,$$

s'il existe un opérateur τ \prec -monotone défini sur P_m tel que

$$\tau^{(1)} \prec \tau \prec \tau^{(2)}.$$

Remarquons que ce théorème reste vrai, en général, pour les opérateurs définis dans un espace partiellement ordonné.

À présent, on considère les opérateurs de transition d'une chaîne de Markov homogène $(X_n)_{n \geq 1}$ d'espace d'état (\mathbb{E}, \mathbb{M}) . Les opérateurs de transition sont décrits par leurs fonctions de transition $p(x, \mathcal{B})$,

$$p(x, \mathcal{B}) = P(X_{n+1} \in \mathcal{B} / X_n = x), \quad x \in \mathbb{E} \text{ et } \mathcal{B} \in \mathbb{E},$$

ou bien, dans le cas où les processus sont à valeurs réelles, par leurs distributions de transition

$$p(x, y) = P(X_{n+1} < y / X_n = x), \quad x, y \in \mathbb{E} \subseteq \mathbb{R}.$$

Maintenant, on donne des conditions sur les fonctions de transition, qui assurent la monotonie ou la comparabilité des opérateurs de transition.

Théoreme 2.3.2.

Les opérateurs de transition $\tau^{(1)}$ et $\tau^{(2)}$ satisfont l'inégalité $\tau^{(1)} \prec \tau^{(2)}$ si et seulement si leurs fonctions $p^{(1)}$ et $p^{(2)}$ satisfont

$$p^{(1)}(x, \cdot) \prec p^{(2)}(x, \cdot), \quad \forall x \in \mathbb{E}.$$

2.3.2 Conditions de monotonie et de comparabilité

Pour l'étude de la monotonie et de la comparabilité des chaînes de Markov homogènes, on peut énoncer les deux théorèmes suivants qui constituent un outil important pour prouver la monotonie interne et/ou externe de ces modèles stochastiques.

Théoreme 2.3.3.

Une chaîne de Markov homogène $(X_n)_{n \geq 1}$, de fonction de transition p , est non décroissante (resp. non croissante) par rapport à l'ordre partiel " \prec " si :

$$X1 \prec X2 \text{ (resp. } X2 \prec X1),$$

et si p est \prec -monotone.

Théoreme 2.3.4.

Deux chaînes de Markov homogènes $(X_n)_{n \geq 1}$ et $(Y_n)_{n \geq 1}$, de fonctions de transition $p^{(1)}$ et $p^{(2)}$ respectivement, satisfont l'inégalité :

$$X_n \prec Y_n, \forall n \in \mathbb{N}^*,$$

si $X_1 \prec X_2$ et s'il existe une fonction de transition p \prec -monotone telle que :

$$p^{(1)}(x, \cdot) \prec p(x, \cdot) \prec p^{(2)}(x, \cdot), \quad \forall x \in \mathbb{E}^*.$$

2.4 Distributions non-paramétriques

Les notions de vieillissement et de relations d'ordre entre variables aléatoires sont étroitement liées. Nous présentons les principaux ordres permettant de comparer des variables aléatoires puis les notions de vieillissement. Cette présentation sera cependant partielle car l'activité scientifique sur ces sujets est importante. Il est donc difficile de prétendre faire un exposé exhaustif. L'un des états de l'art les plus récents sur ce sujet est [8, 15], mais on peut citer aussi [1, 7, 21].

En théorie de fiabilité, les classes de distributions nous renseignent sur la notion de jeunesse ou de vieillesse du système du point de vue de sa durée de vie résiduelle con- naissant son âge (propriété qualitative). La connaissance de la classe (d'âge) de la loi de fiabilité d'un équipement permet une aide à la décision.

Dans cette section, sont présentées aussi les principales classes de distributions de survie recensées dans la littérature de fiabilité ces dernières années.

Les distributions non-paramétriques ont été introduites pour l'étude de certains problèmes en relation avec la théorie de fiabilité. Elles permettent ainsi de modéliser et caractériser des propriétés qualitatives telles que le vieillissement et le rajeunissement du système.

Ces distributions sont utilisées actuellement dans divers domaines de la modélisation stochastique : analyse de survie (médecine), files d'attente, ordonnancement, théorie de décision, économie, gestion des stocks [19].

Définition 2.4.1. Soient X et X_τ des variables aléatoires représentant respectivement la durée de vie et la durée de vie résiduelle d'un élément, et soient F et F_τ leurs distributions respectives. On dit que F est :

- NBU (New Better than Used), si $F \leq_{st} F_\tau$, $(0 < \tau < \infty)$.
- NWU (New Worse than Used), si $F_\tau \leq_{st} F$, $(0 < \tau < \infty)$.
- NBUE (New Better than Used in Expectation), si $E(X_\tau) \leq E(X)$, $(0 < \tau < \infty)$.
- NWUE (New Worse than Used in Expectation), si $E(X) \leq E(X_\tau)$, $(0 < \tau < \infty)$.

Proposition 2.4.1. *Soit la variable aléatoire X de fonction de répartition F ayant une moyenne finie m .*

1. *Si F est NBU (resp. NWU), alors :*

$$F \leq_{st} \exp(\lambda), \quad (\text{resp. } F \geq_{st} \exp(\lambda)).$$

pour un certain $\lambda \leq m^{-1}$ (resp. $\lambda \geq m^{-1}$), avec la possibilité d'avoir une égalité seulement si $F = \exp(m^{-1})$.

2. *Si F est NBUE (resp. NWUE), alors*

$$F \leq_v \exp(m^{-1}), \quad (\text{resp. } F \geq_v \exp(m^{-1})).$$

Conclusion

Dans ce chapitre, on a présenté quelques concepts de base de la théorie des ordres stochastiques et la monotonie des processus stochastiques. Ces techniques seront appliquées, dans le chapitre suivant sur le réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants.

Chapitre 3

Inégalités stochastique pour le réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants

Parmi les approches permettant d'étudier la monotonie du réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants on rencontre celle basée sur la théorie générale des ordres stochastiques.

Dans ce chapitre, on commence par déduire la chaîne de Markov induite d'un réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants, en suite on utilise la théorie générale des ordres stochastiques pour étudier la monotonie de l'opérateur de transition associé à la chaîne de Markov induite de modèle considéré et les conditions de comparabilité des distributions stationnaires.

3.1 Description mathématique du modèle

Considérons un réseau de files d'attente à deux stations en tandem avec blocage et rappels "[M/G/1/1 \rightarrow ./G/1/1] avec rappels constants" illustré dans la figure suivante : Les clients ar-

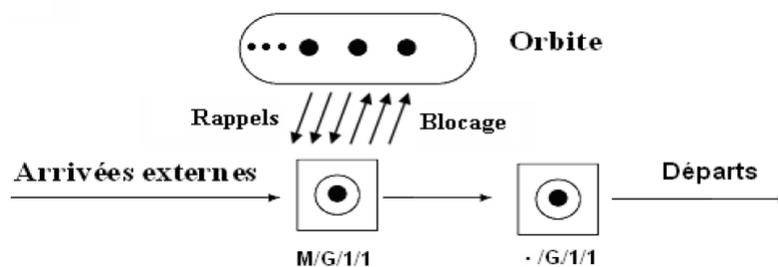


FIGURE 3.1 – Deux stations en tandem avec rappels et blocage

rivent à la première station, selon un processus poissonnien de paramètre λ . Chaque client reçoit une partie de son service à la première station et puis accède à la seconde station pour poursuivre son service. Il n'y a aucune file d'attente intermédiaire, un client dont le service dans la station 1 est accompli ne peut accéder à la station 2 si cette dernière est occupée. Au lieu de cela, le client reste à la station 1 et le système sera bloqué jusqu'à ce que la station 2 se libère. Les clients qui trouvent la station 1 occupée ou bloquée sont obligés de solliciter une autre fois ce système jusqu'à ce que ce dernier soit libre pour les accueillir. Ainsi, ces clients se comporteront comme des rappels, ils ne rejoignent pas une file d'attente mais ils sont placés dans une orbite

de capacité infinie pour réessayer de demander service suivant une politique de rappels dite rappels constants. Suivant cette politique de rappels, le paramètre du temps exponentiel de chaque client dans l'orbite est $\frac{\mu}{n}$, où n est la longueur de la file dans l'orbite. Donc, l'intensité totale des rappels est μ (pour les différentes interprétations de la politique des rappels voir Farahmand [6]). Si le serveur de la station 1 est libre lors d'une tentative de rappel, alors le premier client dans la file de l'orbite reçoit son service immédiatement, autrement il répète sa demande plus tard. Les durées de service dans les deux stations 1 et 2 sont des variables aléatoires indépendantes avec des fonctions de densité $b_i(x)$, des fonctions de répartition $B_i(x)$ et des moyennes $\frac{1}{\mu}$, pour $i = 1, 2$, respectivement.

Soit la variable aléatoire $X(t)$ représentant le nombre de clients dans l'orbite à l'instant t et pour $l = 1, 2$, les variables :

$$\xi^l(t) = \begin{cases} 0 & \text{si le } l^{\text{ème}} \text{ serveur est à vide à l'instant } t, \\ 1 & \text{si le } l^{\text{ème}} \text{ serveur est en activité à l'instant } t, \\ 2 & \text{si le } l^{\text{ème}} \text{ serveur est bloqué à l'instant } t. \end{cases}$$

L'état du réseau $[M/G/1/1 \rightarrow ./G/1/1]$ avec rappels considéré est complètement décrit par le processus $Y(t) = (X(t), \xi^1(t), \xi^2(t))$. On note par d_n , $n \in \mathbb{N}$, l'instant du $n^{\text{ème}}$ départ de la station 1. On suppose, sans perte de généralité, que $d_0 = 0$. Si on note $Y_n = Y(d_n + 0)$, alors il est clair que $Y_n = (X(d_n + 0), \xi^1(d_n + 0), \xi^2(d_n + 0)) = (X_n, 0, 0)$. Donc, le processus $Y(t)$ est semi-génératif avec un processus markovien de renouvellement induit $(X, D) = \{X_n, d_n : n \in \mathbb{N}\}$.

3.2 La chaîne de Markov induite

Le processus induit $(X, D) = \{X_n, d_n : n \in \mathbb{N}\}$ est une chaîne de Markov, on la note par (X_n) . Cette chaîne de Markov $(X_n) = X(d_n + 0)$ induite aux instants de départ du $n^{\text{ème}}$ client, de la première station, est irréductible et apériodique avec la matrice de transition $P = \{p_{ij}\}_{ij}$ définie par :

$$p_{ij} = \begin{cases} \int_0^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} f_0(t) dt, & \text{si } i = 0, \\ \int_0^{\infty} \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t} f_1(t) dt + \int_0^{\infty} \frac{(\lambda t)^{j-i+1}}{(j-i+1)!} e^{-\lambda t} f_2(t) dt, & \text{si } 1 \leq i \leq j + 1, \\ 0. & \text{sinon.} \end{cases}$$

$$\text{avec : } \begin{cases} f_0(t) = \int_0^{\infty} \lambda e^{-\lambda w} \left[\frac{d}{dt} (B_1(t) B_2(t+w)) \right] dw, \\ f_1(t) = \int_0^{\infty} \lambda e^{-(\lambda+\mu)w} \left[\frac{d}{dt} (B_1(t) B_2(t+w)) \right] dw, \\ f_2(t) = \int_0^{\infty} \mu e^{-(\lambda+\mu)w} \left[\frac{d}{dt} (B_1(t) B_2(t+w)) \right] dw. \end{cases}$$

3.3 Notations

Soient deux réseaux de files d'attente $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants Σ_1 et Σ_2 ayant les paramètres suivant :

$\lambda^{(i)}$: taux d'arrivées.

$\mu^{(i)}$: taux de rappels.

$B_1^{(i)}(t), B_2^{(i)}$ distribution du temps de service de la première et la deuxième stations respectivement.

$\pi_n^{(i)}$: la distribution stationnaire du nombre de clients dans le système.

3.4 Monotonie de la chaîne de Markov incluse

Les probabilités de transition en un pas de la chaîne de Markov incluse pour le réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ donnés par la formule suivante :

$$p_{n,m} = \begin{cases} \int_0^{\infty} \left(\frac{(\lambda t)^h}{h!} e^{-\lambda t} f_0(t) \right) dt, & \text{si } n = 0, \\ \int_0^{\infty} \left(\frac{(\lambda t)^{m-n}}{(m-n)!} e^{-\lambda t} f_1(t) \right) dt + \int_0^{\infty} \left(\frac{(\lambda t)^{m-n+1}}{(m-n+1)!} e^{-\lambda t} f_2(t) \right) dt, & \text{si } 1 \leq n \leq m+1, \\ 0. & \text{sinon.} \end{cases}$$

Soit l'opérateur de transition τ de la chaîne de Markov incluse (X_i) qui, pour chaque distribution $p = (p_n)_{n \geq 0}$, on associe une distribution $\tau p = q = (q_m)_{m \geq 0}$ telle que

$$q_m = \sum_{n \geq 0} p_n p_{nm}. \quad (3.1)$$

Les deux Théorèmes suivants nous montrent que l'opérateur de transition τ est monotone par rapport aux ordres stochastique et convexe croissant.

Théoreme 3.4.1.

L'opérateur de transition τ est monotone par rapport à l'ordre stochastique. C'est-à-dire, pour deux distributions quelconques $p^{(1)}$ et $p^{(2)}$, l'inégalité $p^{(1)} \leq_{st} p^{(2)}$ implique $\tau p^{(1)} \leq_{st} \tau p^{(2)}$.

preuve Un opérateur est monotone par rapport à l'ordre stochastique si et seulement si on a l'inégalité suivante :

$$\bar{p}_{n-1,m} \leq \bar{p}_{n,m}, \quad \forall n, m, \quad (3.2)$$

avec,

$$\begin{aligned} \bar{p}_{n,m} &= \sum_{k=m}^{+\infty} p_{n,k} = \sum_{k=m}^{+\infty} \left[\int_0^{\infty} \left(\frac{(\lambda t)^{k-n}}{(k-n)!} e^{-\lambda t} f_1(t) \right) dt + \int_0^{\infty} \left(\frac{(\lambda t)^{k-n+1}}{(k-n+1)!} e^{-\lambda t} f_2(t) \right) dt \right] \\ &= \int_0^{\infty} \left(\sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n}}{(k-n)!} e^{-\lambda t} f_1(t) \right) dt + \int_0^{\infty} \left(\sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} e^{-\lambda t} f_2(t) \right) dt \\ &= \int_0^{\infty} \left(\frac{(\lambda t)^{m-n}}{(m-n)!} e^{-\lambda t} f_1(t) + \sum_{k=m+1}^{+\infty} \frac{(\lambda t)^{k-n}}{(k-n)!} e^{-\lambda t} f_1(t) \right) dt + \int_0^{\infty} \left(\frac{(\lambda t)^{m-n+1}}{(m-n+1)!} e^{-\lambda t} f_2(t) + \sum_{k=m+1}^{+\infty} \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} e^{-\lambda t} f_2(t) \right) dt \\ &= \int_0^{\infty} \left(\frac{(\lambda t)^{m-n}}{(m-n)!} e^{-\lambda t} f_1(t) + \sum_{h=m}^{+\infty} \frac{(\lambda t)^{h-n+1}}{(h-n+1)!} e^{-\lambda t} f_1(t) \right) dt + \int_0^{\infty} \left(\frac{(\lambda t)^{m-n+1}}{(m-n+1)!} e^{-\lambda t} f_2(t) + \sum_{h=m}^{+\infty} \frac{(\lambda t)^{h-n+2}}{(h-n+2)!} e^{-\lambda t} f_2(t) \right) dt. \end{aligned}$$

$$\begin{aligned} \bar{p}_{n-1,m} &= \sum_{k=m}^{+\infty} p_{n-1,k} = \sum_{k=m}^{+\infty} \left[\int_0^{\infty} \left(\frac{(\lambda t)^{k-n+1}}{(k-n+1)!} e^{-\lambda t} f_1(t) \right) dt + \int_0^{\infty} \left(\frac{(\lambda t)^{k-n+2}}{(k-n+2)!} e^{-\lambda t} f_2(t) \right) dt \right] \\ &= \int_0^{\infty} \left(\sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} e^{-\lambda t} f_1(t) \right) dt + \int_0^{\infty} \left(\sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n+2}}{(k-n+2)!} e^{-\lambda t} f_2(t) \right) dt. \end{aligned}$$

$$\begin{aligned}
\bar{p}_{n,m} - \bar{p}_{n-1,m} &= \int_0^\infty \left(\frac{(\lambda t)^{m-n}}{(m-n)!} e^{-\lambda t} f_1(t) + \sum_{h=m}^{+\infty} \frac{(\lambda t)^{h-n+1}}{(h-n+1)!} e^{-\lambda t} f_1(t) - \sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} e^{-\lambda t} f_1(t) \right) dt \\
&+ \int_0^\infty \left(\frac{(\lambda t)^{m-n+1}}{(m-n+1)!} e^{-\lambda t} f_2(t) + \sum_{h=m}^{+\infty} \frac{(\lambda t)^{h-n+2}}{(h-n+2)!} e^{-\lambda t} f_2(t) - \sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n+2}}{(k-n+2)!} e^{-\lambda t} f_2(t) \right) dt \\
&= \int_0^\infty \frac{(\lambda t)^{m-n}}{(m-n)!} e^{-\lambda t} f_1(t) dt + \int_0^\infty \frac{(\lambda t)^{m-n+1}}{(m-n+1)!} e^{-\lambda t} f_2(t) dt \geq 0.
\end{aligned}$$

Par conséquent, l'inégalité (3.12) est vérifiée.

Alors, l'opérateur de transition τ est monotone par rapport à l'ordre stochastique.

Théorème 3.4.2. *L'opérateur de transition τ est monotone par rapport à l'ordre convexe. C'est-à-dire, pour deux distributions quelconques $p^{(1)}$ et $p^{(2)}$, l'inégalité $p^{(1)} \leq_v p^{(2)}$ implique $\tau p^{(1)} \leq_v \tau p^{(2)}$.*

preuve τ est monotone par rapport à l'ordre convexe si et seulement si :

$$2\bar{p}_{n,m} \leq \bar{p}_{n-1,m} + \bar{p}_{n+1,m}, \quad \forall n, m \quad (3.3)$$

où,

$$\begin{aligned}
\bar{p}_{n,m} &= \sum_{k=m}^{+\infty} \bar{p}_{n,k} = \sum_{k=m}^{+\infty} \sum_{l=k}^{+\infty} p_{n,l} \\
&= \sum_{k=m}^{\infty} \sum_{l=k}^{+\infty} \int_0^\infty \left(\frac{(\lambda t)^{l-n}}{(l-n)!} e^{-\lambda t} f_1(t) \right) dt + \sum_{k=m}^{\infty} \sum_{l=k}^{+\infty} \int_0^\infty \left(\frac{(\lambda t)^{l-n+1}}{(l-n+1)!} e^{-\lambda t} f_2(t) \right) dt \\
&= \int_0^\infty \left[\sum_{k=m}^{+\infty} \sum_{l=k}^{+\infty} \frac{(\lambda t)^{l-n}}{(l-n)!} e^{-\lambda t} f_1(t) \right] dt + \int_0^\infty \left[\sum_{k=m}^{+\infty} \sum_{l=k}^{+\infty} \frac{(\lambda t)^{l-n+1}}{(l-n+1)!} e^{-\lambda t} f_2(t) \right] dt. \\
&= \int_0^\infty \left[\sum_{k=m}^{\infty} \left(\frac{(\lambda t)^{k-n}}{(k-n)!} e^{-\lambda t} f_1(t) + \sum_{l=k+1}^{+\infty} \frac{(\lambda t)^{l-n}}{(l-n)!} e^{-\lambda t} f_1(t) \right) \right] dt \\
&+ \int_0^\infty \left[\sum_{k=m}^{\infty} \left(\frac{(\lambda t)^{k-n+1}}{(k-n+1)!} e^{-\lambda t} f_2(t) + \sum_{l=k+1}^{+\infty} \frac{(\lambda t)^{l-n+1}}{(l-n+1)!} e^{-\lambda t} f_2(t) \right) \right] dt \\
&= \int_0^\infty \left[\sum_{k=m}^{\infty} \left(\frac{(\lambda t)^{k-n}}{(k-n)!} e^{-\lambda t} f_1(t) + \sum_{h=k}^{+\infty} \frac{(\lambda t)^{h-n+1}}{(h-n+1)!} e^{-\lambda t} f_1(t) \right) \right] dt \\
&+ \int_0^\infty \left[\sum_{k=m}^{\infty} \left(\frac{(\lambda t)^{k-n+1}}{(k-n+1)!} e^{-\lambda t} f_2(t) + \sum_{h=k}^{+\infty} \frac{(\lambda t)^{h-n+2}}{(h-n+2)!} e^{-\lambda t} f_2(t) \right) \right] dt \\
&= \int_0^\infty \left(\sum_{k=m}^{\infty} \frac{(\lambda t)^{k-n}}{(k-n)!} e^{-\lambda t} f_1(t) + \sum_{k=m}^{\infty} \sum_{h=k}^{+\infty} \frac{(\lambda t)^{h-n+1}}{(h-n+1)!} e^{-\lambda t} f_1(t) \right) dt \\
&+ \int_0^\infty \left(\sum_{k=m}^{\infty} \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} e^{-\lambda t} f_2(t) + \sum_{k=m}^{\infty} \sum_{h=k}^{+\infty} \frac{(\lambda t)^{h-n+2}}{(h-n+2)!} e^{-\lambda t} f_2(t) \right) dt.
\end{aligned}$$

$$\begin{aligned}
\bar{p}_{n+1,m} + \bar{p}_{n-1,m} - 2\bar{p}_{n,m} &= \int_0^\infty \left(\sum_{k=m}^\infty \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} f_1(t) + \sum_{k=m}^\infty \frac{(\lambda t)^{k-n}}{(k-n)!} f_1(t) - 2 \sum_{k=m}^\infty \frac{(\lambda t)^{k-n}}{(k-n)!} f_1(t) \right) e^{-\lambda t} dt \\
&+ \int_0^\infty \left(2 \sum_{k=m}^\infty \sum_{h=k}^{+\infty} \frac{(\lambda t)^{h-n+1}}{(h-n+1)!} f_1(t) - 2 \sum_{k=m}^\infty \sum_{h=k}^{+\infty} \frac{(\lambda t)^{h-n+1}}{(h-n+1)!} f_1(t) \right) e^{-\lambda t} dt \\
&+ \int_0^\infty \left(\sum_{k=m}^\infty \frac{(\lambda t)^{k-n}}{(k-n)!} f_2(t) + \sum_{k=m}^\infty \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} f_2(t) - 2 \sum_{k=m}^\infty \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} f_2(t) \right) e^{-\lambda t} dt \\
&- \int_0^\infty \left(2 \sum_{k=m}^\infty \sum_{h=k}^{+\infty} \frac{(\lambda t)^{h-n+2}}{(h-n+2)!} f_2(t) - 2 \sum_{k=m}^\infty \sum_{h=k}^{+\infty} \frac{(\lambda t)^{h-n+2}}{(h-n+2)!} f_2(t) \right) e^{-\lambda t} dt.
\end{aligned}$$

$$\begin{aligned}
\bar{p}_{n+1,m} + \bar{p}_{n-1,m} - 2\bar{p}_{n,m} &= \int_0^\infty \left(\sum_{k=m}^\infty \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} f_1(t) - \sum_{k=m}^\infty \frac{(\lambda t)^{k-n}}{(k-n)!} f_1(t) \right) e^{-\lambda t} dt \\
&+ \int_0^\infty \left(\sum_{k=m}^\infty \frac{(\lambda t)^{k-n}}{(k-n)!} f_2(t) - \sum_{k=m}^\infty \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} f_2(t) \right) e^{-\lambda t} dt \\
&= \int_0^\infty \left(\frac{(\lambda t)^{m-n-1}}{(m-n-1)!} f_1(t) + \sum_{k=m+1}^\infty \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} f_1(t) - \sum_{k=m}^\infty \frac{(\lambda t)^{k-n}}{(k-n)!} f_1(t) \right) e^{-\lambda t} dt \\
&+ \int_0^\infty \left(\frac{(\lambda t)^{m-n}}{(m-n)!} f_2(t) + \sum_{k=m+1}^\infty \frac{(\lambda t)^{k-n}}{(k-n)!} f_2(t) - \sum_{k=m}^\infty \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} f_2(t) \right) e^{-\lambda t} dt \\
&= \int_0^\infty \left(\frac{(\lambda t)^{m-n-1}}{(m-n-1)!} f_1(t) + \sum_{h=m}^\infty \frac{(\lambda t)^{h-n}}{(h-n)!} f_1(t) - \sum_{k=m}^\infty \frac{(\lambda t)^{k-n}}{(k-n)!} f_1(t) \right) e^{-\lambda t} dt \\
&+ \int_0^\infty \left(\frac{(\lambda t)^{m-n}}{(m-n)!} f_2(t) + \sum_{h=m}^\infty \frac{(\lambda t)^{h-n-1}}{(h-n-1)!} f_2(t) - \sum_{k=m}^\infty \frac{(\lambda t)^{k-n+1}}{(k-n+1)!} f_2(t) \right) e^{-\lambda t} dt.
\end{aligned}$$

$$\bar{p}_{n+1,m} + \bar{p}_{n-1,m} - 2\bar{p}_{n,m} = \int_0^\infty \left(\frac{(\lambda t)^{m-n-1}}{(m-n-1)!} e^{-\lambda t} f_1(t) \right) dt + \int_0^\infty \left(\frac{(\lambda t)^{m-n}}{(m-n)!} e^{-\lambda t} f_2(t) \right) dt \geq 0.$$

Par conséquent, l'inégalité (3.13) est vérifiée.

Alors, l'opérateur de transition τ est monotone par rapport à l'ordre convexe.

3.5 Comparaison des opérateurs de transition

Dans cette section, on compare les opérateurs de transition associés à deux chaînes de Markov incluse ayant la même structure mais de paramètres différents.

Théoreme 3.5.1. *Si $\lambda^{(1)} \leq \lambda^{(2)}$, $\mu^{(1)} \leq \mu^{(2)}$, $B_1^{(1)} \leq_{st} B_1^{(2)}$, $B_2^{(1)} \leq_{st} B_2^{(2)}$ alors $\tau^{(1)} \leq \tau^{(2)}$. C'est-à-dire, pour n'importe quelle distribution p on a $\tau^{(1)}p \leq_{st} \tau^{(2)}p$.*

Preuve. *D'après le Théorème 2.3.2, nous devons vérifier les inégalités suivantes pour l'ordre stochastique.*

$$\bar{p}_{n,m}^{(1)} \leq \bar{p}_{n,m}^{(2)}, \quad \forall 1 \leq n \leq m+1$$

Ce qui revient à montrer :

$$\begin{aligned}
\bar{p}_{n,m}^{(1)} &= \sum_{k=m}^{\infty} \int_0^{\infty} \frac{(\lambda^{(1)}t)^{k-n}}{(k-n)!} e^{-\lambda^{(1)}t} f_1^{(1)}(t) dt + \sum_{k=m}^{\infty} \int_0^{\infty} \frac{(\lambda^{(1)}t)^{k-n+1}}{(k-n+1)!} e^{-\lambda^{(1)}t} f_2^{(1)}(t) dt \\
&\leq \sum_{k=m}^{\infty} \int_0^{\infty} \frac{(\lambda^{(2)}t)^{k-n}}{(k-n)!} e^{-\lambda^{(2)}t} f_1^{(2)}(t) dt + \sum_{k=m}^{\infty} \int_0^{\infty} \frac{(\lambda^{(2)}t)^{k-n+1}}{(k-n+1)!} e^{-\lambda^{(2)}t} f_2^{(2)}(t) dt = \bar{p}_{n,m}^{(2)} \\
\bar{p}_{n,m}^{(1)} &= \int_0^{\infty} \left(\sum_{k=m}^{\infty} \frac{(\lambda^{(1)}t)^{k-n}}{(k-n)!} e^{-\lambda^{(1)}t} f_1^{(1)}(t) \right) dt + \int_0^{\infty} \left(\sum_{k=m}^{\infty} \frac{(\lambda^{(1)}t)^{k-n+1}}{(k-n+1)!} e^{-\lambda^{(1)}t} f_2^{(1)}(t) \right) dt \\
&\leq \int_0^{\infty} \left(\sum_{k=m}^{\infty} \frac{(\lambda^{(2)}t)^{k-n}}{(k-n)!} e^{-\lambda^{(2)}t} f_1^{(2)}(t) \right) dt + \int_0^{\infty} \left(\sum_{k=m}^{\infty} \frac{(\lambda^{(2)}t)^{k-n+1}}{(k-n+1)!} e^{-\lambda^{(2)}t} f_2^{(2)}(t) \right) dt = \bar{p}_{n,m}^{(2)}. \quad (3.4)
\end{aligned}$$

on considère la fonction

$$g_m(t, \lambda) = \sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n}}{(k-n)!} e^{-\lambda t}.$$

En prenant sa dérivée par rapport à t , on obtient

$$\begin{aligned}
\frac{\partial}{\partial t} g_m(t, \lambda) &= \sum_{k=m}^{+\infty} \left[-\lambda e^{-\lambda t} \frac{(\lambda t)^{k-n}}{(k-n)!} + (k-n) \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n)!} \right] \\
&= -\sum_{k=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n}}{(k-n)!} + \sum_{k=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} \\
&= -\sum_{k=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n}}{(k-n)!} + \lambda e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} + \sum_{k=m+1}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} \\
&= -\sum_{k=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n}}{(k-n)!} + \lambda e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} + \sum_{h=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{h-n}}{(h-n)!} \\
&= \lambda e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} > 0.
\end{aligned}$$

donc $g_n(t, \lambda)$ est une fonction croissante en t .

La dérivée par rapport à λ est :

$$\begin{aligned}
\frac{\partial}{\partial \lambda} g_m(t, \lambda) &= \frac{\partial}{\partial \lambda} \left[\sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n}}{(k-n)!} e^{-\lambda t} \right] \\
&= \sum_{k=m}^{+\infty} \left[-t e^{-\lambda t} \frac{(\lambda t)^{k-n}}{(k-n)!} + (k-n) t e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n)!} \right] \\
&= -\sum_{k=m}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n}}{(k-n)!} + \sum_{k=m}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} \\
&= -\sum_{m=n}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n}}{(k-n)!} + t e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} + \sum_{k=m+1}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} \\
&= -\sum_{m=n}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n}}{(k-n)!} + t e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} + \sum_{h=m}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{h-n}}{(h-n)!} \\
&= t e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} > 0. \quad (3.5)
\end{aligned}$$

puisque la fonction $g_m(t, \lambda)$ est monotone par rapport à λ et que $\lambda^{(1)} \leq \lambda^{(2)}$, alors

$$g_m(t, \lambda^{(1)}) = \sum_{k=m}^{\infty} \frac{(\lambda^{(1)}t)^{k-n}}{(k-n)!} e^{-\lambda^{(1)}t} \leq \sum_{k=m}^{\infty} \frac{(\lambda^{(2)}t)^{k-n}}{(k-n)!} e^{-\lambda^{(2)}t} = g_m(t, \lambda^{(2)}).$$

et pour

$$f_1^{(1)}(t) \leq f_1^{(2)}(t).$$

alors :

$$g_m(t, \lambda^{(1)})f_1^{(1)}(t) = \sum_{k=m}^{\infty} \frac{(\lambda^{(1)}t)^{k-n}}{(k-n)!} e^{-\lambda^{(1)}t} f_1^{(1)}(t) \leq \sum_{k=m}^{\infty} \frac{(\lambda^{(2)}t)^{k-n}}{(k-n)!} e^{-\lambda^{(2)}t} f_1^{(2)}(t) = g_m(t, \lambda^{(2)})f_1^{(2)}(t). \quad (3.6)$$

et on considère la fonction

$$h_m(t, \lambda) = \sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} e^{-\lambda t}.$$

En prenant sa dérivée par rapport à t , on obtient

$$\begin{aligned} \frac{\partial}{\partial t} h_m(t, \lambda) &= \sum_{k=m}^{+\infty} \left[-\lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} + (k-n-1) \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-2}}{(k-n-1)!} \right] \\ &= -\sum_{k=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} + \sum_{k=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-2}}{(k-n-2)!} \\ &= -\sum_{k=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} + \lambda e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} + \sum_{k=m+1}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-2}}{(k-n-2)!} \\ &= -\sum_{k=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} + \lambda e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} + \sum_{h=m}^{+\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{h-n-1}}{(h-n-1)!} \\ &= \lambda e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} > 0. \end{aligned}$$

donc $h_m(t, \lambda)$ est une fonction croissante en t .

La dérivée par rapport à λ est :

$$\begin{aligned} \frac{\partial}{\partial \lambda} h_m(t, \lambda) &= \frac{\partial}{\partial \lambda} \left[\sum_{k=m}^{+\infty} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} e^{-\lambda t} \right] \\ &= \sum_{k=m}^{+\infty} \left[-t e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} + (k-n-1) t e^{-\lambda t} \frac{(\lambda t)^{k-n-2}}{(k-n-1)!} \right] \\ &= -\sum_{k=m}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} + \sum_{k=m}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n-2}}{(k-n-2)!} \\ &= -\sum_{m=n}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} + t e^{-\lambda t} \frac{(\lambda t)^{m-n-2}}{(m-n-2)!} + \sum_{k=m+1}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n-2}}{(k-n-2)!} \\ &= -\sum_{m=n}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{k-n-1}}{(k-n-1)!} + t e^{-\lambda t} \frac{(\lambda t)^{m-n-2}}{(m-n-2)!} + \sum_{h=m}^{+\infty} t e^{-\lambda t} \frac{(\lambda t)^{h-n-1}}{(h-n-1)!} \\ &= t e^{-\lambda t} \frac{(\lambda t)^{m-n-1}}{(m-n-1)!} > 0. \end{aligned} \quad (3.7)$$

puisque la fonction $h_m(t, \lambda)$ est monotone par rapport à λ et que $\lambda^{(1)} \leq \lambda^{(2)}$, alors

$$h_m(t, \lambda^{(1)}) = \sum_{k=m}^{\infty} \frac{(\lambda^{(1)}t)^{k-n-1}}{(k-n-1)!} e^{-\lambda^{(1)}t} \leq \sum_{k=m}^{\infty} \frac{(\lambda^{(2)}t)^{k-n-1}}{(k-n-1)!} e^{-\lambda^{(2)}t} = h_m(t, \lambda^{(2)}).$$

et pour

$$f_2^{(1)}(t) \leq f_2^{(2)}(t).$$

alors :

$$h_m(t, \lambda^{(1)})f_2^{(1)}(t) = \sum_{k=m}^{\infty} \frac{(\lambda^{(1)}t)^{k-n-1}}{(k-n-1)!} e^{-\lambda^{(1)}t} f_2^{(1)}(t) \leq \sum_{k=m}^{\infty} \frac{(\lambda^{(2)}t)^{k-n-1}}{(k-n-1)!} e^{-\lambda^{(2)}t} f_2^{(2)}(t) = h_m(t, \lambda^{(2)})f_2^{(2)}(t). \quad (3.8)$$

Par conséquent, des inégalités (3.6) et (3.8), on aura :

$$g_m(t, \lambda^{(1)})f_1^{(1)}(t) + h_m(t, \lambda^{(1)})f_2^{(1)}(t) \leq g_m(t, \lambda^{(2)})f_1^{(2)}(t) + h_m(t, \lambda^{(2)})f_2^{(2)}(t). \quad (3.9)$$

On intègre terme à terme (3.9) sur $[0, +\infty]$ par rapport à t , on obtient :

$$\int_0^{+\infty} \left(g_m(t, \lambda^{(1)})f_1^{(1)}(t) + h_m(t, \lambda^{(1)})f_2^{(1)}(t) \right) dt \leq \int_0^{+\infty} \left(g_m(t, \lambda^{(2)})f_1^{(2)}(t) + h_m(t, \lambda^{(2)})f_2^{(2)}(t) \right) dt.$$

C'est-à-dire,

$$\int_0^{+\infty} \left(\sum_{k=m}^{\infty} \frac{(\lambda^{(1)}t)^{k-n}}{(k-n)!} f_1^{(1)}(t) + \sum_{k=m}^{\infty} \frac{(\lambda^{(1)}t)^{k-n-1}}{(k-n-1)!} f_2^{(1)}(t) \right) e^{-\lambda^{(1)}t} dt \leq \int_0^{+\infty} \left(\sum_{k=m}^{\infty} \frac{(\lambda^{(2)}t)^{k-n}}{(k-n)!} f_1^{(2)}(t) + \sum_{k=m}^{\infty} \frac{(\lambda^{(2)}t)^{k-n-1}}{(k-n-1)!} f_2^{(2)}(t) \right) e^{-\lambda^{(2)}t} dt.$$

donc, l'inégalité (3.4) est vérifiée.

Théoreme 3.5.2. Si $\lambda^{(1)} \leq \lambda^{(2)}$, $\mu^{(1)} \leq \mu^{(2)}$, $B_1^{(1)} \leq_v B_1^{(2)}$, $B_2^{(1)} \leq_v B_2^{(2)}$ alors $\tau^{(1)} \leq \tau^{(2)}$. C'est-à-dire, pour n'importe quelle distribution p on a $\tau^{(1)}p \leq_v \tau^{(2)}p$.

Preuve. D'après le Théorème 2.3.2, nous devons vérifier les inégalités suivantes pour l'ordre convexe.

$$\bar{p}_{n,m}^{(1)} \leq \bar{p}_{n,m}^{(2)}, \quad \forall 1 \leq n \leq m+1$$

Ce qui revient à montrer :

$$\begin{aligned} \bar{p}_{n,m}^{(1)} &= \sum_{k=m}^{\infty} \sum_{l=k}^{\infty} \int_0^{\infty} \frac{(\lambda^{(1)}t)^{l-n}}{(l-n)!} e^{-\lambda^{(1)}t} f_1^{(1)}(t) dt + \sum_{k=m}^{\infty} \sum_{l=k}^{\infty} \int_0^{\infty} \frac{(\lambda^{(1)}t)^{l-n+1}}{(l-n+1)!} e^{-\lambda^{(1)}t} f_2^{(1)}(t) dt \\ &\leq \sum_{k=m}^{\infty} \sum_{l=k}^{\infty} \int_0^{\infty} \frac{(\lambda^{(2)}t)^{l-n}}{(l-n)!} e^{-\lambda^{(2)}t} f_1^{(2)}(t) dt + \sum_{k=m}^{\infty} \sum_{l=k}^{\infty} \int_0^{\infty} \frac{(\lambda^{(2)}t)^{l-n+1}}{(l-n+1)!} e^{-\lambda^{(2)}t} f_2^{(2)}(t) dt = \bar{p}_{n,m}^{(2)} \\ \bar{p}_{n,m}^{(1)} &= \int_0^{\infty} \sum_{k=m}^{\infty} \left[\sum_{l=k}^{\infty} \frac{(\lambda^{(1)}t)^{l-n}}{(l-n)!} e^{-\lambda^{(1)}t} f_1^{(1)}(t) \right] dt + \int_0^{\infty} \sum_{k=m}^{\infty} \left[\sum_{l=k}^{\infty} \frac{(\lambda^{(1)}t)^{l-n+1}}{(l-n+1)!} e^{-\lambda^{(1)}t} f_2^{(1)}(t) \right] dt \\ &\leq \int_0^{\infty} \sum_{k=m}^{\infty} \left[\sum_{l=k}^{\infty} \frac{(\lambda^{(2)}t)^{l-n}}{(l-n)!} e^{-\lambda^{(2)}t} f_1^{(2)}(t) \right] dt + \int_0^{\infty} \sum_{k=m}^{\infty} \left[\sum_{l=k}^{\infty} \frac{(\lambda^{(2)}t)^{l-n+1}}{(l-n+1)!} e^{-\lambda^{(2)}t} f_2^{(2)}(t) \right] dt = \bar{p}_{n,m}^{(2)} \end{aligned}$$

$$\begin{aligned}
\bar{p}_{n,m}^{(1)} &= \int_0^\infty \left(\sum_{k=m}^\infty g_k(t, \lambda^{(1)}) f_1^{(1)}(t) \right) dt + \int_0^\infty \left(\sum_{k=m}^\infty h_k(t, \lambda^{(1)}) f_2^{(1)}(t) \right) dt \\
&\leq \int_0^\infty \left(\sum_{k=m}^\infty g_k(t, \lambda^{(2)}) f_1^{(2)}(t) \right) dt + \int_0^\infty \left(\sum_{k=m}^\infty h_k(t, \lambda^{(2)}) f_2^{(2)}(t) \right) dt = \bar{p}_{n,m}^{(2)}. \quad (3.10)
\end{aligned}$$

avec,

$$g_k(t, \lambda^{(i)}) = \sum_{l=k}^{+\infty} \frac{(\lambda^{(i)} t)^{l-n}}{(l-n)!} e^{-\lambda^{(i)} t}, \quad h_k(t, \lambda^{(i)}) = \sum_{l=k}^{+\infty} \frac{(\lambda^{(i)} t)^{l-n-1}}{(l-n-1)!} e^{-\lambda^{(i)} t}, \quad i = 1, 2.$$

Les fonctions $g_k(t, \lambda)$ sont croissantes par rapport à λ , alors la fonction définie par :

$$\bar{g}_m(t, \lambda) = \sum_{k=m}^\infty g_k(t, \lambda) \quad \text{l'est aussi. et que } \lambda^{(1)} \leq \lambda^{(2)}, \text{ alors :}$$

$$\bar{g}_m(t, \lambda^{(1)}) \leq \bar{g}_m(t, \lambda^{(2)})$$

et pour

$$f_1^{(1)}(t) \leq f_1^{(2)}(t).$$

alors

$$\bar{g}_m(t, \lambda^{(1)}) f_1^{(1)}(t) \leq \bar{g}_m(t, \lambda^{(2)}) f_1^{(2)}(t). \quad (3.11)$$

Les fonctions $h_m(t, \lambda)$ sont croissantes par rapport à λ , alors la fonction définie par :

$$\bar{h}_m(t, \lambda) = \sum_{k=m}^\infty h_k(t, \lambda) \quad \text{l'est aussi. et que } \lambda^{(1)} \leq \lambda^{(2)}, \text{ alors :}$$

$$\bar{h}_m(t, \lambda^{(1)}) \leq \bar{h}_m(t, \lambda^{(2)}).$$

et pour

$$f_2^{(1)}(t) \leq f_2^{(2)}(t).$$

alors

$$\bar{h}_m(t, \lambda^{(1)}) f_2^{(1)}(t) \leq \bar{h}_m(t, \lambda^{(2)}) f_2^{(2)}(t). \quad (3.12)$$

Par conséquent, des inégalités (3.11) et (3.12), on aura :

$$\bar{g}_m(t, \lambda^{(1)}) f_1^{(1)}(t) + \bar{h}_m(t, \lambda^{(1)}) f_2^{(1)}(t) \leq \bar{g}_m(t, \lambda^{(2)}) f_1^{(2)}(t) + \bar{h}_m(t, \lambda^{(2)}) f_2^{(2)}(t). \quad (3.13)$$

On intègre terme à terme (3.13) sur $[0, +\infty]$ par rapport à t , on obtient :

$$\int_0^\infty \left(\bar{g}_m(t, \lambda^{(1)}) f_1^{(1)}(t) \right) dt + \int_0^\infty \left(\bar{h}_m(t, \lambda^{(1)}) f_2^{(1)}(t) \right) dt \leq \int_0^\infty \left(\bar{g}_m(t, \lambda^{(2)}) f_1^{(2)}(t) \right) dt + \int_0^\infty \left(\bar{h}_m(t, \lambda^{(2)}) f_2^{(2)}(t) \right) dt.$$

Finalement, l'inégalité (3.10) est vérifiée.

3.6 Inégalités stochastiques des distributions stationnaires du nombre de clients dans le système

Les deux théorèmes suivants donnent les conditions de comparabilité des distributions stationnaires du nombre de clients, pour deux systèmes de files d'attente $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants, par rapport aux ordres partiels : stochastiques et convexe.

Théoreme 3.6.1. On considère Σ_1 et Σ_2 deux réseaux de files d'attente $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants ayant les paramètres $\lambda^{(1)}, \mu^{(1)}, B_1^{(1)}(t), B_2^{(1)}(t)$ et $\lambda^{(2)}, \mu^{(2)}, B_1^{(2)}(t), B_2^{(2)}(t)$ respectivement, et soient $\pi_n^{(1)}, \pi_n^{(2)}$ les distributions stationnaires du nombre de clients dans chaque système, alors si les inégalités suivantes ont lieu :

$$\lambda^{(1)} \leq \lambda^{(2)}, \mu^{(1)} \leq \mu^{(2)}, B_1^{(1)} \leq_{so} B_1^{(2)}, B_2^{(1)} \leq_{so} B_2^{(2)},$$

on a aussi les inégalités suivantes sur les distributions stationnaires

$$\{\pi_n^{(1)}\} \leq_{so} \{\pi_n^{(2)}\}, \quad \text{où } so = st \text{ (ou } v).$$

Preuve. D'après le Théorèmes (3.5.2), les inégalités $\lambda^{(1)} \leq \lambda^{(2)}, \mu^{(1)} \leq \mu^{(2)}, B_1^{(1)} \leq_{so} B_1^{(2)}, B_2^{(1)} \leq_{so} B_2^{(2)}$ alors $\tau^{(1)} \leq \tau^{(2)}$. C'est-à-dire, pour n'importe quelle distribution p on a

$$\tau^{(1)}p \leq_{so} \tau^{(2)}p \tag{3.14}$$

Des Théorèmes (3.5.1) l'opérateur $\tau^{(2)}$ associé à la chaîne de Markov incluse, du deuxième système, est monotone. C'est-à-dire, pour deux distributions quelconques $p_1^{(2)}, p_2^{(2)}$ telles que $p_1^{(2)} \leq_{so} p_2^{(2)}$, on a

$$\tau^{(2)}p_1^{(2)} \leq_{so} \tau^{(2)}p_2^{(2)}. \tag{3.15}$$

Cependant, de l'inégalité (3.14), on obtient

$$\tau^{(1)}p^{(1)} \leq_{so} \tau^{(2)}p^{(1)}. \tag{3.16}$$

Il existe une probabilité $p_1^{(2)}$ telle qu'on ait l'inégalité suivante

$$\tau^{(2)}p^{(1)} \leq_{so} \tau^{(2)}p_1^{(2)}. \tag{3.17}$$

En combinant les inégalités (3.15)–(3.17), on obtient le résultat suivant

$$\tau^{(1)}p^{(1)} \leq_{so} \tau^{(2)}p^{(2)}. \tag{3.18}$$

pour deux distributions quelconques $p^{(1)}, p^{(2)}$.

L'inégalité (3.18) peut être réécrite de la manière suivante

$$\tau^{(1)}p_n^{(1)} = P\left(Z_k^{(1)} = n\right) \leq_{so} P\left(Z_k^{(2)} = n\right) = \tau^{(2)}p_n^{(2)}. \tag{3.19}$$

Quand $k \rightarrow \infty$, on déduit le résultat recherché, c'est-à-dire, $\{\pi_n^{(1)}\} \leq_{so} \{\pi_n^{(2)}\}$.

Théoreme 3.6.2. Si pour le réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants, la distribution de temps de service est NBUE (New Better than Used Expectation), (respectivement NWUE-New Worse than Used in Expectation), alors la distribution stationnaire du nombre de clients dans ce système est inférieure (respectivement supérieure), par rapport à l'ordre convexe, à la distribution stationnaire du nombre de clients dans le réseaux $\{M/M/1/1 \rightarrow ./M/1/1\}$ avec rappels constants.

Preuve. Considérons un système de files d'attente $\{M/M/1/1 \rightarrow ./M/1/1\}$ avec rappels constants, du serveur avec les mêmes paramètres : taux d'arrivée λ , taux de rappels μ , temps moyen de service

β_1^1 et β_2^1 , que le système $\{M/M/1/1 \rightarrow ./M/1/1\}$ avec rappels constants, mais avec un temps de service exponentiellement distribué avec de taux $\theta_1 = \frac{1}{\beta_1^1}$ et $\theta_2 = \frac{1}{\beta_2^1}$.

$$B_1^* = \begin{cases} 1 - e^{-\frac{x}{\beta_1^1}}, & \text{si } x \geq 0, \\ 0, & \text{si } x < 0. \end{cases}$$

$$B_2^* = \begin{cases} 1 - e^{-\frac{x}{\beta_2^1}}, & \text{si } x \geq 0, \\ 0, & \text{si } x < 0. \end{cases}$$

D'après la proposition (2.4.1), Si $B_k(x)$, $k=1,2$ est NBUE (respectivement NWUE), alors

$$B_k(x) \leq_v B_k^*(x) \text{ (respectivement } B_k(x) \geq_v B_k^*(x)).$$

alors d'après le Théorème (3.6.1) on déduit que la distribution du nombre de clients dans le système $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants, est inférieure (respectivement supérieure) la la distribution du nombre de clients dans le système $\{M/M/1/1 \rightarrow ./M/1/1\}$ avec rappels constants.

Conclusion

Dans ce chapitre, on a montré la monotonie de l'opérateur de transition de la chaîne de Markov incluse par rapport à l'ordre stochastique et convexe. On a aussi obtenu des conditions de comparabilité des deux opérateurs de transition. En outre, nous avons prouvé que la distribution stationnaire du nombre de clients liée au modèle considéré, est majorée (respectivement minorée) par la distribution stationnaire du nombre de clients dans un réseaux $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants, si la distribution des temps de service est NBUE (respectivement NWUE).

Conclusion générale

Ce mémoire est consacré a la départ de la théorie de files d'attente, Nous nous sommes intéressés aux modèles $M/G/1$ avec rappels.

Nous avons aussi étudié quelques problèmes de comparabilité pour l'analyse du système $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants, en utilisant la méthode de comparaison stochastique. L'avantage de ce type de méthodes d'approximation réside dans le fait que des résultats explicites puissent être obtenus pour des situations relativement complexes. Ceci nous a permis d'obtenir les conditions qui assurent la monotonie de l'opérateur de transition associé à la chaîne de Markov induite.

Certaines caractéristiques de ce modèle sont en fait connues, mais ici, on considère deux problèmes :

1. La monotonie de la chaîne de Markov incluse.
2. La comparabilité des modèles de ce types, mais ayant des paramètres différents. Concernant le premier problème, il s'avère que la chaîne de Markov incluse n'est pas monotone en général. Ceci nous a permis d'obtenir les conditions qui assurent la monotonie de l'opérateur de transition associé à la chaîne de Markov incluse. Pour le second problème, nous avons établi des conditions sous lesquelles les opérateurs de transition ainsi que les distributions stationnaires de deux chaînes de Markov incluse associées à deux systèmes $\{M/G/1/1 \rightarrow ./G/1/1\}$ avec rappels constants, ayant la même structure mais avec des paramètres différents, sont comparables au sens des ordres stochastique et convexe.

Parmi les perspectives de recherche, nous citons :

- Appliquer la méthode de comparaison stochastique pour étudier les propriétés de monotonie du notre modèle relativement à l'ordre stochastique et concave.

Bibliographie

- [1] Boualem. M, Djellab. N and Aïssani. D (2009). Stochastic inequalities for M/G/1 retrial queues with vacations and constant retrial policy, International Journal MCM (Mathematical and Computer Modelling), Doi : 10.1016/j. mcm. 2009.03.009, ElsevierEd, 50.207-212.
- [2] Boualem. M and Djellab. N and Assani. D (2012). Stochastic approximations and monotonicity of a single server feedback retrial queue, Mathematical Problems in Engineering, (2012), Article ID 536982, 1-13.
- [3] Boualem. M, Djellab. N and Assani. D (2014). Stochastic bounds for a single server queue with general retrial times. Bulletin of the Iranian Mathematical Society, 40(1), 183-198.
- [4] Dietrich Stoyan (1983). Comparaison Methods for Queues and Other Stochastic Models, New York.
- [5] Djouhra, Dad (2011). Modélisation et simulation du flux dans un réseau pour la régulation du trafic, Magister Dissertation, Université Ahmed Ben Bella d’Oran 1 Es Senia, Algeria.
- [6] Faramand .K (1990). Single Line Queue with Repeated Demands. Queue. Syst., 6 :223–228.
- [7] Genedenko. B, Beliaev. Y and Soloviev .A (1972), Méthodes Mathématiques en théorie de la fiabilité. Chapman and Hall, London chains, 745. Wiley-Interscience, 605 Third Avenue, NewYork, NY 10158-0012.
- [8] Gine. E, Houdre. C and Nualart. D(2003), Stochastic inequalities and applications, Progress in Probability 56, Birkhauser
- [9] Gunter Bolch, S. GREINER, H.de MEER, and K.S. TRIVEDI (1998). Queuing Network and Markov
- [10] H. KAMEDA and Y. ZHANG (1995). Uniqueness of the solution for optimal static routing in open bcmp queueing networks. Mathematical and computer modelling, 22(10-12), pp.119-130.
- [11] Ivo Adan and Jacques Resing (26 March 2015). Queueing Systems, Département of Mathematics and Computing Science Eindhoven University of Technology, 182 pages.
- [12] Jean-Michel Fourneau, Nihal Pekergin. Dominique Verchère (1998). Réseaux généralisés multiclassés avec synchronisation cycliques. RAIRO.recherche opérationnelle, tome 32.
- [13] Kebbas Salah (2013). Contribution à la Correction et l’amélioration de la Qualité de Service dans une Entreprise Publique, en utilisant les Réseaux de Files d’attente, MÉMOIRE DE MAGISTER, 100 pages.
- [14] Khalil. Z and Falin. G (1994). Stochastic inequalities for M/G/1 retrial queues, Oper. Res. Lett, 16,285-290.
- [15] Krishna Kumar. B, Pavai Madheswari . S and Anantha Lakshmi. S. R (2013), An M/G/1 Bernoulli feedback retrial queueing system with negative customers Oper Res Int J 13 :187-210 DOI 10.1007/s12351-011-0107-5
- [16] Lala Maghnia Alem, Mohamed Boualem and Djamil Aissani. Stochastic comparison bounds for an $M_1, M_2/G_1, G_2/1$ retrial queue with two way communication. Hacettepe Journal of Mathematics and Statistics, 48(4) (2019), 1185-1200.

- [17] Lala Maghnia Alem, Mohamed Boualem and Djamil Aissani. Bounds of the stationary distribution in M/G/1 retrial queue with two way communication and n type of outgoing call. Yugoslav Journal of Operations Research, 29(3) (2019), 375-391.
- [18] Mohamed Jarraya. Cours modélisation des réseaux et évaluation des performances. 2006
- [19] Müller. A and Stoyan. D(2002). Comparison methods for stochastic models and risk, John Wiley and Sons, LTD.
- [20] Oukid. N and Aissani. A (2009). Bounds on busy period for queues with breakdowns, Advances and Applications in Statistics, 11, 137-156.
- [21] S. FDIDA, et G. HERBUTERNE, Méthodes exactes d'analyse de performance des réseaux Edition Lavoisier. 2004. opérationnelle université A/Mira de Bejaia.
- [22] T. Yang and J. G. C. Templeton. A survey on retrial queues. Queueing Systems, 2 : 203-233,, 1987.

Résumé

En raison de la complexité des modèles d'attente $\{M/G/1/1 \rightarrow ./G/1/1\}$ Avec rappels constant, les résultats analytiques sont généralement difficiles à obtenir ou ne sont pas très exploitables en pratique. Pour résoudre le problème, on a fait appel à la méthode de comparaison stochastique pour étudier les propriétés de monotonie et de comparabilité du réseau.

En outre, on a montré la monotonie de l'opérateur de transition associé à la chaîne de Markov induite, ainsi qu'on a prouvé que la distribution stationnaire du nombre de clients liée au modèle considéré, est majorée (respectivement minorée) par la distribution stationnaire du nombre de clients dans un réseau $[M/M/1/1 \rightarrow ./M/1/1]$ avec rappels constants, si la distribution des temps de service est NBUE (respectivement NWUE).

Mots-clés : réseaux de files d'attente avec rappels, Chaîne de Markov induite, comparaison stochastique, monotonie.

Abstract

Due to the complexity of expectations models $\{M/G/1/1 \rightarrow ./G/1/1\}$ With constant reminders, the analytical results are generally difficult to obtain or are not very exploitable in practice. To solve the problem, the stochastic comparison methods were used to study the properties of monotony and comparability of the network.

Moreover, we have shown the monotony of the transition operator associated with the induced Markov chain, as it has been proved that the stationary distribution of the number of clients related to the model considered, is increased (respectively minus) by the stationary distribution of the number of clients in a network $[M/M/1/1 \rightarrow ./M/1/1]$ with constant callbacks, if the distribution of service times is NBUE (respectively NWUE).

Keywords : retrial networks, embedded Markov chain, stochastic comparison, monotonicity.