République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2019/2020

Système de file d'attente Markovien avec vacances simple

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université Dr Tahar Moulay - Saïda
Discipline : MATHÉMATIQUES
Spécialité : ASSPA

Par

KEDDARI RABIE ABDEL-MOHCEN

Sous la direction de

Dr. M.KADI

Soutenu le 16 septembre 2020 devant le jury composé de

Dr. F.MOKHTARI	Université Dr Tahar Moulay - Saïda	Présidente
Dr. M.KADI	Université Dr Tahar Moulay - Saïda	Encadreur
Dr. N.AIT OUALI	Université Dr Tahar Moulay - Saïda	Examinatrice
Dr. N.HACHEMI	Université Dr Tahar Moulay - Saïda	Examinatrice

Table des matières

Remerciement

Je remercie vivement mon encadreur Mr. M.KADI, docteur à l'université de saida de m'avoir proposé le sujet de ce mémoire, pour sa patience et ses conseils qui m'ont été d'un grand apport, et encouragement pendant toute la durée de l'élaboration de ce travail. Je tiens également à remercier :

- Le président de jury, d'avoir eu l'aimabilité d'accepter la présidence du jury de soutenance.
- Les membres de jury pour leur participation et leur dévouement.

≺Merci à vous tous≻

Dédicace

Je dédie ce modeste travail à mes parents qui m'ont aidé pour arriver à ce que je suis aujourd'hui.

A mon très cher ami RAMDANI MOHAMED pour tout son aide, son soutien et ses encouragements.

A mes frères MOHAMED et YASSIN à qui je souhaite une longue vie pleine de bonheur et réussite.

Introduction générale

Les évolutions récentes de l'environnement technologique dues à l'accroissement des besoins liés à l'industrialisation font des modéles de files d'attente l'un des outils indispensables dans la modélisation mathématique. En effet, une grande classe de modèles mathématiques provient de la théorie de files d'attente dont les bases présentées dans les ouvrages de L.Kleinrock s'ouvrent à de nombreuses applications. Un modèle typique de files d'attente nécessite la définition des processus d'inter arrivées et de durée de service des clients, la taille de la file qui peut être finie ou non, ainsi que la discipline de service. Tous ces paramètres sont indiqués dans la notation dite de Kendall . Depuis la première étude probabiliste réalisée en 1917 par l'ingénieur Danois Erlang pour modéliser le central téléphonique de Copenhague, la modélisation et l'évaluation des performances des systèmes complexes par les modèles de files d'attente n'ont cessé d'être un domaine de recherche d'un intérêt croissant. Quoique bon nombre de formules magnifiques furent élaborées et suggérées comme étant des solutions analytiques de certains types de problèmes, les travaux effectués dans le domaine des files d'attente restent jusque là limité en raison de l'inexploitabilité pratique de la plupart des résultats connus. La théorie est assez développée pour l'étude d'une file simple.

Ce mémoire est composé de trois chapitres : Dans le premier chapitre, nous présentons les processus stochastiques qui sont les notions de bases de l'étude des systèmes de files d'attente. Ensuite dans le deuxième chapitre, nous introduisons certaines définitions et notations sur la théorie des files d'attente comme (Notation de Kendall, la loi de Little ,...etc)

Et enfin, dans le troisième chapitre on s'intéresse aux systèmes de files d'attente avec vacance simple (unique), nous analysons le modèle $M/M/1$ avec un temps de patience réparti de manière exponentielle, dans le cadre de la politique de vacances simples. Nous dérivons les équations d'équilibre et nous obtenons les équations differentielles pour les PGF des probabilités en régime permanent. Cela nous permet de calculer certaines mesures de performance telles que la taille moyenne des systèmes lorsque le serveur est soit en vacances, soit occupé, la proportion de clients servis et le temps de séjour moyen.

Chapitre 1

Processus stochastiques

1.1 Introduction

Un processus stochastique est une suite de variables aléatoires indexée par $\mathbb{T}$ à valeur dans un ensemble $\mathbb{X}$ qui premet à décriver un phénomène aléatoire en fonction du temps. Dans ce chapitre on va présenté plusiers processus stochastiques utilisés pour modéliser des files d'attenats parmis ces processus le procusse de poisson qui est très important dans les phénomènes de comptage puis le processus de renouvlement, et enfin le processus de naisance et mort qui pèermet de modéliser les processus stochastiques évoluant dans le temps.

Définition 1.1.1. ***(Processus stochastique)***

Un processus stochastique est une famille $(\mathbf{X}(t), t \in \mathbb{T})$ *de variables aléatoires définie sur un même espace de probabilité.*
Généralement $\mathbb{X}(t)$ *représente l'état du processus stochastique au temps* t.

- *Si* $\mathbb{T}$ *est dans* $[0; \infty)$ *alors le processus stochastique est dit un processus à temps continu.*
- *Si* $\mathbb{T}$ *est dénombrable, i.e.* $\mathbb{T} \subseteq \mathbb{N}$ *alors nous disons que* $\mathrm{X(t)_{t \in T}}$ *est un processus à temps discret.*

Avec $\mathrm{X}(t)$ *représente l'état du processus stochastique au temps* t *.*

Remarque 1.1.1. *L'ensemble* $\mathbb{T}$ *est muni d'une relation d'ordre* $\leq$*, qui est de plus totale : étant donnés* $(s, t) \in \mathbb{T}^2$ *on a* $s \leq t$ *ou* $t \leq s$*. On peut aussi considérer des processus sur un horizon de temps fini*

- *Dans le cas discret, on considère* $\mathbb{T} = \{0, \ldots, N\}$*, pour un certain instant final* N
- *Dans le cas continu, on pose* $\mathbb{T} = [0, T]$.

1.2 Processus de comptage

Définition 1.2.1. *(processus de comptage)*

Soit $N(t)$ un processus stochastique et si $N(t)$ représente le nombre total des événements qui sont arrivés avant l'instant t,on dit que $N(t)$ est un processus de comptage discret à temps continu.
Tout processus de comptage vérifie les propriétés suivantes :

1. *Le nombre $N(t)$ est à valeurs entières positives,pour tout $t \geq 0$*
2. *La fonction $t \mapsto N(t)$ est croissante.*
3. *La différence $N(t) - N(s)$ représente le nombre d'évènement se produisant dans l'intervalle de temps $]s, t]$,Pour tout couple $(s, t)(0 < s < t)$*

Le processus des temps d'inter-arrivées $\{W_n, n \in N_0\}$ ou $\forall\, n \in N_0$ la variable aléatoire W_n est le temps d'attente entre les $(n-1)^{ieme}, n^{ieme}$ occurrences, est un processus peut étre associé au processus des temps d'occurrence c-à-d :

$$W_n = T_n - T_{n-1}$$

avec T_n est le temps d'arrivé du n^{ieme} client.

Demonstration : On a W_n= $T_n - T_{n-1}$

$$\begin{aligned} W_1 + W_2 + ... + W_n &= T_1 - T_0 + T_2 - T_1 + T_3 - T_2 + ... + T_{n-1} - T_{n-2} + T_n - T_{n-1} \\ &= T_0 + T_n \\ &= T_n \quad car \quad T_0 = 0 \end{aligned}$$

1.3 Processus de renouvellement

Les processus de Poisson et de renouvellement sont des processus aléatoires de comptage à temps continu qui conviennent à la description des phénomènes dont les occurrences surviennent en des temps successifs aléatoires. Ils sont utiles à la modélisation des files d'attente apparaissant à l'entrée de services et dans les réseaux de communication, et permettent de résoudre des problèmes de maintenance par exemple.

Définition 1.3.1. *(processus de renouvellement)*

Le processus de renouvellement est un processus de comptage pour le quel les temps entre deux arrivés consecutives sont des variables aleatoires i.i.d. Les temps de renouvellement (ou les temps de la n-iême arrivée) sont :

$$A_n = \sum_{i=1}^{n} a_i, n = 0, 1, 2, \dots$$

On voit que le nombre d'arrivées avant le temps t, i.e. le processus

$$(N_t)_{t \in \mathbb{R}_+} = \sup_k \{k : A_k \leq t\}$$

est un processus de comptage.

1.3.1 Processus à accroissements indépendants

Définition 1.3.2. ***(processus à accroissements indépendants)***

Un processus stochastique $\mathbb{X}_t$ tel que $\mathbb{X}_0 = 0$ est dit un processus à accroissements indépendants si pour tout suite finie $0 < t_1 < t_2 < t_3 \cdots < t_n$ les variables aléatoires $\mathbb{X}_{t_1}\mathbb{X}_{t_2} - \mathbb{X}_{t_1}, \cdots, \mathbb{X}_{t_n} - \mathbb{X}_{t_{n-1}}$ sont indépendantes.

Un processus à accroissements stationnaires est un processus stochastique qui vérifie la propriété

– *La loi de $\mathbb{X}_{t+s} - \mathbb{X}_t$ ne dépend pas de t pour tout $t > 0$*

1.4 Processus de Poisson

Le processus de Poisson sert â modéliser l'occurrence d'évènements successifs. Chaque éévènement est tel que dans un intervalle de temps $(t, t + \Delta t)$ avec Δt petit.

Il peut servir à modéliser par exemple :

– Les appels téléphoniques arrivant dans une centrale.
– Les temps d'arrivée de clients à une caisse.
– Passage d'un autobus.
– Arrivée d'un client â un guichet,

Définition 1.4.1. ***(processus de poisson)***

Un processus de comptage $(N_t)_{t \in R}$ tel que $N_0 = 0$ est dit processus de Poisson si

* $(N_t)_{t \in R}$ *est stationnaire,*
* $(N_t)_{t \in R}$ *est un processus à accroissements indépendants stationnaire .,*
* *il existe $\lambda > 0$ tel que, pour tout $t \geq 0$, la variable aléatoire N_t suive la loi de Poisson de paramètre λt.*

1.4.1 Loi de Poisson et la loi exponentielle

Définition 1.4.2. ***(Loi de poisson)***

Une variable aléatoire X de loi Poisson de paramètre $\lambda > 0$ est de valeur dans $\mathbb{N}$ de probabilité

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda k}$$

avec λ le paramètre de la loi

Propriété 1.4.1.1. *Soit* X *une variable aléatoire de loi de Poisson $\mathcal{P}(\mu)$*

* *La fonction génératrice des moments est*

$$\varphi(t) = \mathbb{E}\left[e^{tX}\right] = e^{\lambda\left(e^t - 1\right)}$$

* *La moyenne et la variance sont*

$$\mathbb{E}[X] = \lambda \quad \mathrm{Var}(X) = \lambda$$

preuve :

$$\begin{aligned}\varphi(t) &= \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^t\right)^k}{k!} \\ &= e^{-\lambda} e^{\lambda e^t}\end{aligned}$$

calculons les dérivées de la fonction génératrice

$$\varphi'(t) = \lambda e^t e^{\lambda\left(e^t - 1\right)}$$
$$\varphi''(t) = \left(1 + \lambda e^t\right) \lambda e^{\lambda} e^{\lambda\left(e^t - 1\right)}$$

D'où

$$\mathbb{E}[X] = \varphi'(0) = \lambda$$
$$\mathrm{Var}(X) = \varphi''(0) - \mathbb{E}[X]^2 = \lambda$$

Définition 1.4.3. ***(Loi exponentielle)***

Soit la variable aléatoire $\mathbf{X}$ *continue de loi exponentielle de paramètre $\mu > 0$ à valeurs positives de densité*

$$f(x) = \mu e^{-\mu x} \mathbb{1}_{x \geqslant 0}$$

Propriété 1.4.1.2. *La loi exponentielle de paramètre μ est notée $\mathcal{E}(\mu)$.*

Soit X une variable aléatoire de loi $\mathcal{E}(\mu)$.

- Sa fonction de répartition est : $F(t) = \begin{cases} 1-e^{-\mu t} & si \quad t \geq 0 \\ 0 & si \quad t < 0. \end{cases}$

- Sa fonction génératrice des moments est : $\varphi(t) = \mathbb{E}\left[e^{tX}\right] = \begin{cases} \infty & si \quad t \geqslant \mu \\ \dfrac{\mu}{\mu - t} & si \quad t < \mu \end{cases}$

-Sa moyenne et sa variance sont $\mathbb{E}[X] = \dfrac{1}{\mu} \quad \mathrm{Var}(X) = \dfrac{1}{\mu^2}$

Preuve

Soit X une variable aléatoire de loi $\mathcal{E}(\mu)$

- $F(t)| = \mathbb{P}(X \leqslant t) = 0$ si $t < 0$ car X est une veriable positive et si $t \geqslant 0$ on a

$$\begin{aligned} F(t) &= \int_0^t f_X(x)dx \\ &= \int_0^t \mu e^{-\mu x}dx \\ &= 1 - e^{-\mu t} \end{aligned}$$

* Sa fonction génératrice des moments vérifie

$$\begin{aligned} \varphi(t) &= \mathbb{E}\left[e^{tX}\right] \\ &= \int_0^{+\infty} \mu e^{tx}e^{-\mu x}dx \\ &= \int_0^{+\infty} \mu e^{(t-\mu)x}dx \\ &= \left[\frac{\mu}{t-\mu}e^{(t-\mu)x}\right]_0^{+\infty} \end{aligned}$$

* D'où le résultat.Calculons les dérivées de la fonction génératrice, on a

$$\varphi'(t) = \frac{\mu}{(\mu - t)^2} \quad \varphi''(t) = \frac{2\mu}{(\mu - t)^3}$$

D'où

$$\mathbb{E}[X] = \varphi'(0) = \frac{1}{\mu} \quad \text{et} \quad \mathrm{Var}(X) = \varphi''(0) - \mathbb{E}[X]^2 = \frac{1}{\mu^2}$$

1.4.1.1 Relation entre les distributions Exponentielle et Poissonniene

On a la densité de la loi exponentielle $f(t) = \mu e^{-\mu t}$ Supposons τ est exponentielle d'une espérance $\frac{1}{\mu}$ avec n est de Poisson de moyenne λ on a :

$$\begin{aligned} P(\tau > t) &= 1 - F(t) \\ &= e^{-\mu t} \\ &= P(n = 0 \quad \text{en} \quad t) \\ &= P(0, t) \end{aligned}$$

on note $P(n,t)$ la probabilité d'avoir n unité dans le temps t.

$P(0,t) = e^{-\lambda t}$

$P(1,t) = \int_{\tau=0}^{t} P(0,\tau) f(1-\tau) d\tau = \mu t e^{-\mu t}$

$P(2,t) = \int_{\tau=0}^{t} P(1,\tau) f(1-\tau) d\tau = (\mu t)^2 e^{-\mu t}/2!$

.

.

.

$P(n,t) = \int_{\tau=0}^{t} P(n-1,\tau) f(1-\tau) d\tau = (\mu t)^n e^{-\mu t}/n!$

1.4.1.2 Propriètè sans mémoire de la distribution exponentielle

Une variable aléatoire X est dite sans mémoire (ou sans usure) si :

$\forall x, y \geq 0$

$$\mathrm{P}(X > x + y / X > y) = \mathrm{P}(X > x)$$

Si X modélise la durée de vie d'un individu A, la propriété que X est sans mémoire exprime que A ne vieillit pas

si A a vécu y années, la probabilité pour qu'il vive encore x années est la même que la probabilité pour qu'un individu similaire à A qui vient de naître vive lui aussi x années.

On prouve qu'une variable aléatoire X positive absolument continue est sans mémoire si elle suit une loi exponentielle.

1.4.1.3 L'utilisation de la loi de poisson et exponentielle

Une loi exponentielle modélise la durée de vie d'un phénomène sans mémoire, ou sans vieillissement, ou sans usure : la probabilité que le phénomène dure au moins $s+t$ heures sachant qu'il a déjà duré t heures sera la même que la probabilité de durer s heures à partir de sa mise en fonction initiale. En d'autres termes, le fait que le phénomène ait duré pendant t heures ne change rien à son espérance de vie à partir du temps t.

Un domaine privilégié de la loi exponentielle est le domaine de la radioactivité. Chaque atome radioactif possède une durée de vie qui suit une loi exponentielle. Le paramètre μ s'appelle alors la constante de désintégration.

- La durée de vie moyenne $\frac{1}{\mu}$ s'appelle le temps caractéristique.
- La loi des grands nombres permet de dire que la concentration d'atomes radioactifs va suivre la même loi. La médiane $\frac{\ln(2)}{\mu}$ correspond au temps T nécessaire pour que la population passe à 1/2 de sa population initiale et s'appelle la demi-vie ou période.

En théorie des probabilités et en statistiques, la loi de Poisson est une loi de probabilité discrète qui décrit le comportement du nombre d'événements se produisant dans un intervalle de temps fixé, si ces événements se produisent avec une fréquence moyenne ou espérance connue, et indépendamment du temps écoulé depuis l'événement précédent.

Le domaine d'application de la loi de Poisson a été longtemps limité à celui des événements rares comme les suicides d'enfants, les arrivées de bateaux dans un port ou les accidents dus aux coups de pied de cheval dans les armées.

Mais depuis quelques décennies son champ d'application s'est considérablement élargi. Actuellement, on l'utilise beaucoup dans les télécommunications (pour compter le nombre de communications dans un intervalle de temps donné), la biologie (mutations dans l'expérience de Luria et Delbrück, nombre de potentiels d'actions émis par un neurone en neurosciences), la météorologie, la finance pour modéliser la probabilité de défaut d'un crédit, le Yield Management (American Airlines, Lufthansa pour estimer la demande de passagers), etc.

1.5 Processus de naissance et de mort

Les processus de naissance et de mort sont des cas particuliers de processus de Markov en temps continu où les transitions d'état sont de deux types seulement : les «naissances» où l'état passe de n à n+1 et les morts où l'état passe de n à n-1. Ces processus ont de nombreuses applications en dynamique des populations et dans la théorie des files d'attente

On va modèliser ce processus de telle façon que :

- Le processus est spécifié par les taux de naissance $(\lambda_n)_{n=0\ldots\infty}$ et les taux de mortalité $(\mu_n)_{n=0\ldots\infty}$.
- Il existe une transition vers un état voisin, soit par l'arrivée ou le départ d'un client (naissance et mort) On suppose que $\lambda_0 = 0$.

Si $\pi_n(t)$ est la probabilité de trouver le système dans l'état n avec $(n = 0, 1, 2, ...)$ à l'instant t, alors l'équation de Kolomogorov [13] s 'écrit, pour $n > 0$

$$\pi_n(t + dt) = (1 - (\mu_n + \lambda_n)\, dt)\, \pi_n(t) + \lambda_{n+1}\pi_{n+1}(t)dt + \mu_{n-1}\pi_{n-1}(t)dt + o(dt)$$

On va tendre dt vers 0, pour $n > 0$:

$$\frac{d}{dt}\pi_n(t) = -\left(\mu_n + \lambda_n\right)\pi_n(t) + \lambda_{n+1}\pi_{n+1}(t) + \mu_{n-1}\pi_{n-1}(t)$$

on obtient pour $n = 0$.

$$\frac{d}{dt}\pi_0(t) = -\mu_0\pi_0(t) + \lambda_1\pi_1(t).$$

Un cas particulier du processus de naissance et de mort c'est Le processus de poisson avec $\lambda_n = 0$ et $\mu_n = \mu$ dans ce cas on trouve pas un régime stasionnaire [5] les équations différentielles s'écrivent alors

$$\pi_0(t) = e^{-\mu t} \qquad \frac{d}{dt}\pi_0(t) = -\mu_0\pi_0(t), \qquad \frac{d}{dt}\pi_n(t) = -\mu\left(\pi_n(t) - \pi_{n-1}(t)\right)$$

la solution est $\quad \pi_n(t) = \dfrac{(\mu t)^n e^{-\mu t}}{n!}$

1.5.1 Représentation de transition d'un processus de naissance et mort

- λ_n : taux de naissances quand le nombre de population égale à n.
- μ_n : taux de morts quand le nombre de population égale à n.

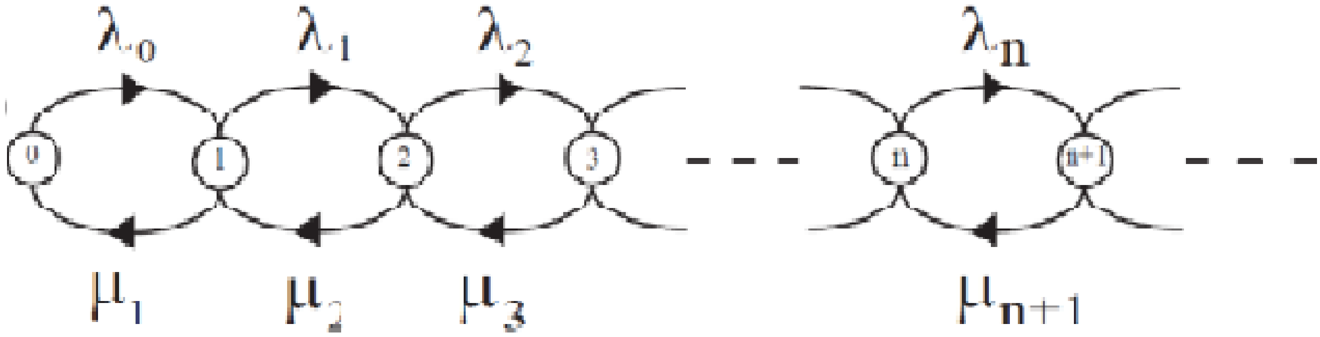

FIGURE 1.1 – Diagramme de transition d'un processus de naissance et de mort

Chapitre 2

Système de files d'attente

La théorie des files d'attente est une théorie mathématique relevant du domaine des probabilités, qui étudie les solutions optimales de gestion des files d'attente, ou queues. Une queue est nécessaire et se créera d'elle-même si ce n'est pas anticipé, dans tous les cas où l'offre est inférieure à la demande, même temporairement. Ce domaine de recherches, né en 1917, des travaux de l'ingénieur danois Erlang sur la gestion des réseaux téléphoniques de Copenhague à partir de 1908, étudie notamment les systèmes d'arrivée dans une queue, les différentes priorités de chaque nouvel arrivant, ainsi que la modélisation statistique des temps d'exécution.
Elle peut s'appliquer à différentes situations :

1. gestion des avions au décollage ou à l'atterrissage.
2. attente des clients et des administrés aux guichets de bétonnières se charger en ciment, ...etc).
3. stockage des programmes informatiques avant leur traitement... etc. .

2.1 Définitions

File d'attente : Quand un client occupe le serveur pendant un certain temps, les autres clients doivent attendre avant d'être servis , formant un ensemble qui s'apelle une file d'attente.
Système d'attente : l'ensemble des clients qui font la queue, avec celui qui se fait servir.
Types de files d'attente intérêt dans les situations où il existe plusieurs stations et plusieurs files d'attente.

1. Files séparées : une file par guichet (par exemple, aux caisses des supermarchés) ; ce système a l'inconvénient de générer des frustrations lorsque certaines files sont plus rapides que d'autres, ou lorsqu'un guichet supplémentaire s'ouvre, permettant aux derniers de passer les premiers.

2. File distribuée ou mutualisée : une seule file alimente plusieurs guichets, ce qui a pour effet d'éviter les inconvénients des files d'attente uniques. Ce type de file réduit le temps d'attente moyen, équilibre le travail des agents/caissiers, garantit un meilleur service à la clientèle et prévient la création d'un goulot causé par un événement spécial ou inattendu survenu dans l'un des guichets.
3. File virtuelle : une prise de ticket permet de conserver l'ordre d'arrivée, sans avoir à faire la queue physiquement ; par exemple, les personnes peuvent s'asseoir en attendant leur tour.
4. File virtuelle mobile : les nouvelles technologies permettent maintenant de prendre rang par internet ou par téléphone, et d'être prévenu par SMS lorsque son tour approche, le temps d'attente ne nécessitant plus une présence physique.
5. File prioritaire : des files plus rapides peuvent être créées, par exemple pour les personnes ayant un handicap, ou pour les personnes ayant une carte de fidélité ; parfois, des files prioritaires payantes peuvent être proposées.

Classification des Systèmes d'attente [10] Pour décrire une file d'attente, on doit donc se donner les éléments suivants :

- La nature du processus des arrivées (flux d'entrée) qui est définie par la distribution des intervalles séparant deux arrivées consécutives.
- La distribution du temps aléatoire de service.
- Le nombre des stations de service montées en parallèle.
- La capacité totale du système N qui représente le nombre maximum de clients pouvant être présents dans le système.

Notations et symboles [2]

- λ : le nombre moyen d'arrivées ;(taux d'arrivée).
- $\frac{1}{\lambda}$: temps moyen séparant deux arrivées consécutives
- μ : le nombre moyen de clients servis (taux de service)
- $\frac{1}{\mu}$: la durée moyenne de service (un client)
- $\overline{N} = \mathbb{E}(N)$: nombre moyen de clients dans le système,
- $\overline{N}_S$: nombre moyen de clients en train d'être servis,
- $\overline{N}_Q$: nombre moyen de clients dans la file d'attente, N_Q, N_S et N sont les v.a. correspondantes,
- $\overline{T}$: temps moyen qu'un client passe dans le système,

- $\overline{T}_S$: temps moyen de service,
- $\overline{T}_Q$: temps moyen d'attente d'un client dans la file. T_Q, T_S et T sont les v.a. correspondantes.

2.2 Description d'une file d'attente simple

Définition 2.2.1. *Une file d'attente est un système caractérisé par un espace d'attente qui contient une ou plusieurs places, et un espace de service composé d'un ou plusieurs serveurs. Les clients arrivent de l'extérieur à des instants aléatoires, ils attendent que l'un des serveurs soit libre pour pouvoir être servi puis quittent le système.*

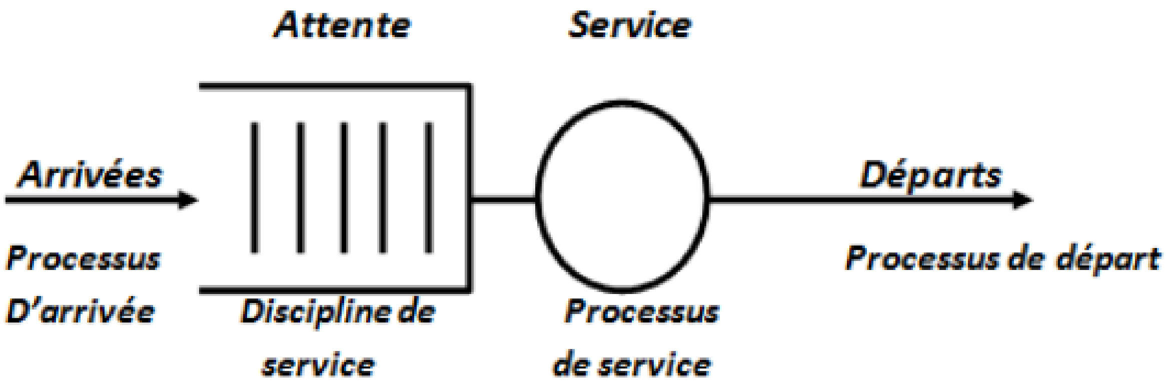

FIGURE 2.1 – le système de file d'attente

2.2.1 Processus d'arrivé :

L'arrivée des clients à la station sera décrite à l'aide d'un processus stochastique de comptage $(N_t)_{t\geq 0}$

Si A_n désigne la variable aléatoire mesurant l'instant d'arrivée du n ième client dans le système, on aura ainsi :$A_0 = 0$ et $A_n = \inf\{t; N_t = n\}$

Si T_n désigne la variable aléatoire mesurant le temps séparant l'arrivée du $(n-1)$ ième et du n ième client [9], on a alors :

$$T_n = A_n - A_{n-1}$$

2.2.2 Processus de service :

Considérons tout d'abord une file à serveur unique. On note Dn la variable aléatoire mesurant l'instant de départ du n^{ime} client du système et Yn la variable aléatoire mesurant le temps de service du n^{ime} client (le temps séparant le début et la fin du service). Un instant de départ correspond toujours à une fin de service, mais ne correspond pas forcément à un début de service . Il se peut en effet qu'un client qui quitte la station laisse celle-ci vide. Le serveur est alors inoccupé jusqu'à l'arrivée du prochain client. On note μ le taux de service :

$\frac{1}{\mu}$ est la durée moyenne de service.

2.2.3 Structure et discipline de la file

2.2.3.1 Nombre de serveurs :

Une station peut disposer de plusieurs serveurs en parallèle. Soit C le nombre de serveurs. Dès qu'un client arrive à la station, soit il y a un serveur libre, le client entre instantanément en service, soit tous les serveurs sont occupés et le client se place dans la file en attente de libération d'un des serveurs. Mais en suppose à la plupart du temps que les serveurs sont identiques et indépendants les uns des autres. Une station particulière est la station IS (infinité servers) dans la quelle le nombre de serveurs est infini. Donc cette station ne comporte pas de file d'attente

2.2.4 Capacité de la file :

La capacité de la file à accueillir des clients en attente de service peut être finie ou infinie. Soit N la capacité de la file, une file à capacité illimitée vérifie $N = \infty$

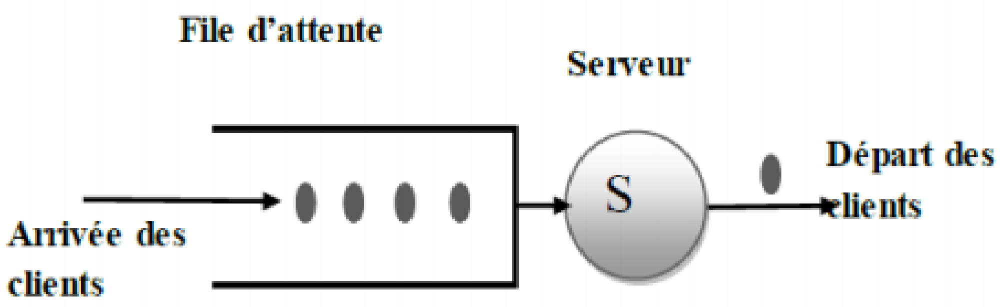

FIGURE 2.2 – Système de file d'attente avec un serveur unique.

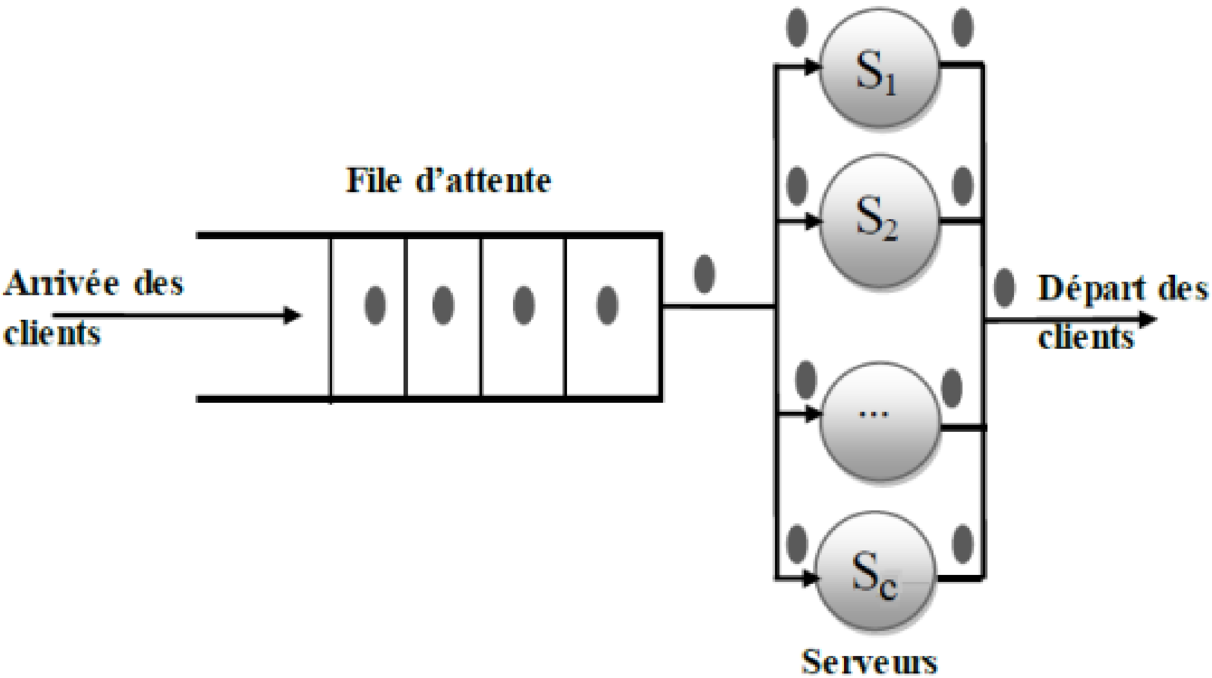

FIGURE 2.3 – Système de file d'attente à C serveurs parallèles.

2.3 Notation de Kendall

La notation suivante, appelée la notation de Kendall [6], est largement utilisée pour classer les différents systèmes de files d'attent :

$$A/Y/C/H/n/Z$$

1. A : indique le processus d'arrivée des clients. Les codes utilisés sont :
   - M (Markov) : inter-arrivées des clients sont indépendamment, identiquement distribuées selon une loi exponentielle. Il correspond à un processus de Poisson ponctuel (propriété sans mémoire),
   - D (Répartition déterministe) : les temps inter-arrivées des clients ou les temps de service sont constants et toujours les mêmes,
   - GI (général indépendant) : Les inter-arrivées des clients ont une distribution générale (il n'y a aucune hypothèse sur la distribution mais les interarrivées sont indépendentes et identiquement distribuées),
   - G (général) : Inter-arrivées de clients ont une distribution générale et peuvent être dépendantes,
   - E_k : Ce symbole désigne un processus où les intervalles de temps entre deux arrivées successives sont des variables aléatoires indépendantes et identiquement

distribuées suivant une loi d'Erlang d'ordre k.

2. Y : décrit la distribution des temps de service d'un client. les codes sont les mêmes que A,
3. C : nombre de serveurs,
4. H : capacité de la file (c'est le nombre de places dans le système en d'autre tèrme c'est le nombre maximal de clients permis dans le système B compris ceux en service,
5. n : population des usagers,
6. z : discipline de service (c'est la façon dont les clients sont ordonnés pour être servi.

Discipline de service :

- FIFO (first in, first out) ou FCFS (first come first served) : c'est la file standard dans laquelle les clients sont servis dans leur ordre d'arrivée. Notons que les disciplines FIFO et FCFS ne sont pas équivalentes lorsque la file contient plusieurs serveurs. Dans la première, le premier client arrivé sera le premier à quitter la file alors que dans la deuxième, il sera le premier à commencer son service. Rien n'empêche alors qu'un client qui commence son service après lui,
- LIFO (last in, first out) ou LCFS (last come, first served). Cela correspond à une pile, dans laquelle le dernier client arrivé (donc posé sur la pile) sera le premier traité (retiré de la pile). A nouveau, les disciplines LIFO et LCFS ne sont équivalentes que pour une file monoserveur
- SIRO (Served In Random Order), les clients sont servis aléatoirement.
- PNPN (Priority service), les clients sont servis selon leur priorité. Tous les clients de la plus haute priorité sont servis premiers, puis les clients de priorité inférieur sont servis, et ainsi de suite.
- PS (Processor Sharing), les clients sont servis de manière égale. La capacité du système est partagée entre les clients.

Remarque : Dans sa version courte, seuls les trois premiers symboles A/Y/C sont utilisés. dans un tel cas, on suppose que la file est régie par une discipline FIFO et que le nombre de places d'attente ainsi que celui des clients susceptibles d'accéder au système sont illimités.

2.4 Loi de Little

La loi de Little est une relation très générale qui s'applique à une grande classe de systèmes. Elle ne concerne que le régime permanent du systéme. Aucune hypothèse sur les variables aléatoires qui caractérisent le système (temps d'inter-arrivées, temps de service,...etc). La seule condition d'application de la loi de Little est que le système soit stable. Le débit du système est alors indifféremment soit le débit d'entrée, soit le débit de sortie :$d_s = d_e = d$ La loi de Little s'exprime telle que dans la propriété suivante :

Théorème 2.4.1. ***(Formule de Little)*** *: Le nombre moyen de clients , le temps moyen passé dans le système et le débit moyen d'un système stable en régime permanent se relient de la façon suivante :*

$$\overline{N} = \lambda_e \overline{T}$$

pour une file$(M/M/1)$ $\lambda_e = \lambda$

On a vu que la loi de Little nous dit qu'il existe une relation entre le nombre moyen de clients dans la file (en attente ou en service) et le temps moyen total de séjour d'un client dans la file(temps d'attente +temps de service)

Remarque 2.4.1. *La loi de Little s'applique à tous les modèles de file d'attente rencontrés en pratique (pas seulement à la file $M/M/1$).*

2.5 Mesure de performance d'une file d'attente

Le but d'étudie une file d'attente ou un réseau de files d'attente c'est d'estimer les performances d'un système dans des conditions de fonctionnement données, et les mesures les plus fréquemment, de manière générale : si le nombre moyen d'arrivées de clients par unité de temps, noté λ, est inférieur au nombre moyen de clients pouvant être servis par unité de temps alors la file est stable.
Si chaque serveur peut traiter λ clients par unité de temps et si le nombre de serveurs est c, une file est stable si et seulement si

$$\lambda < c\mu \Leftrightarrow \rho = \lambda/c\mu < 1$$

où, ρ est appelé l'intensité du trafic.

2.6 File $M/M/1$

La file d'attente M/M/1 est un modèle caractérisé par des arrivées suivants un processus de Poisson de taux λ, temps de service exponentiel de paramètre μ et un seul serveur. Les clients arrivent à la station selon un processus de Poisson de taux λ, si le serveur est vide le client est pris en charge immédiatement sinon il rejoint la file d'attente de capacité illimitée et discipline FIFO.
La file peut être considérée comme un processus de naissance et de mort pour lequel

$$\lambda_n = \lambda \quad \forall n \geq 0$$

$$\mu_n = \begin{cases} \mu & \forall n \geq 1 \\ 0 & \text{si } n = 0 \end{cases}$$

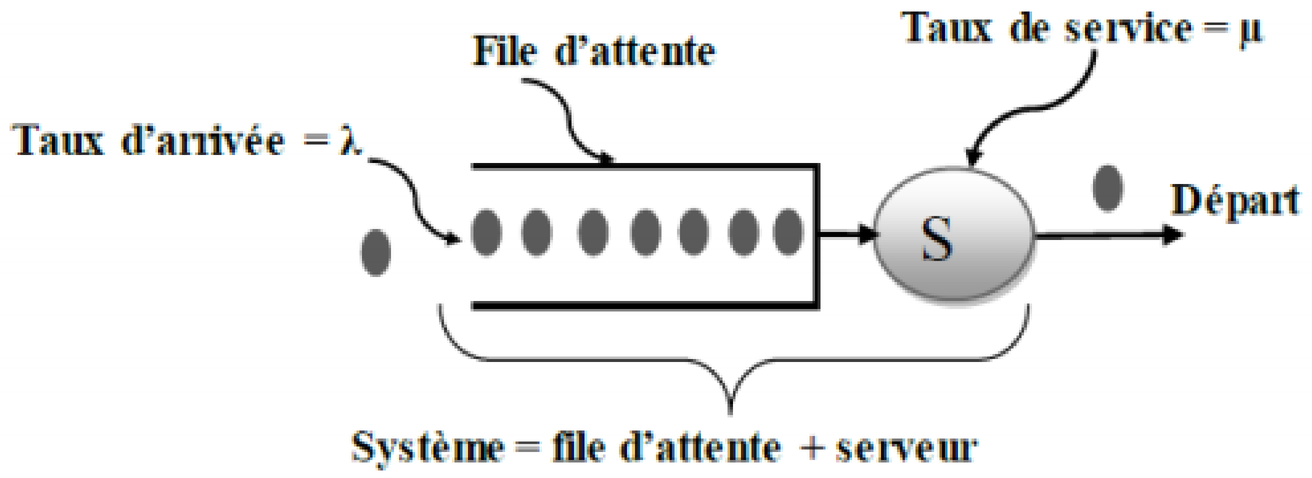

FIGURE 2.4 – La file $M/M/1$

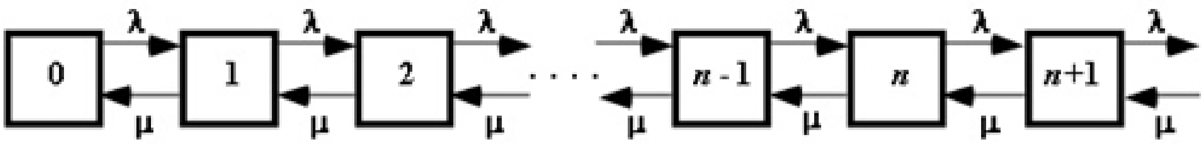

FIGURE 2.5 – Diagramme de transition de la file d'attente $M/M/1$.

Les probabilités d'état $p_n(t) = \mathbb{P}[N(t) = n]$ peuvent être calculées par les équations différentielles de Kolmogorov ci-dessous, connaissant les conditions initiales du processus.

$$p'_n(t) = -(\lambda + \mu)p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t)$$

$$\text{et} \quad p_0(t) = \lambda p_0(t) + \mu p_1(t)$$

Sous l'hypothèse que $\lambda < \mu$ (le taux d'arrivée est plus petit que le taux de service). On a :

$$\rho = \frac{\lambda}{\mu} < 1$$

Les probabilités d'état pour un régime stationnaire du processus sont données par :

$$\pi_n = \pi_0 \rho^n$$
$$\pi_0 = \frac{1}{\sum_{n=0}^{\infty} \rho^n} = 1 - \rho$$

Donc

$$\pi_n = (1-\rho)\rho^n, \quad \forall n \in \mathbb{N}$$

$\pi = \{\pi_n\}_{n \geq 0}$ est appelé distribution stationnaire, elle suit une loi géométrique.(la probabilité stationnaire d'avoir n clients dans la file).

Caractéristiques du système :

– Le nombre moyen de clients dans le système est :

$$\begin{aligned} \overline{N} &= \mathbb{E}(N) \\ &= \sum_{n=1} n\pi_n \\ &= (1-\rho)\sum_{n=1} n\rho^n \end{aligned}$$

D'où :

$$\overline{N} = \frac{\rho}{1-\rho}$$

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

$$\overline{N}_S = 1 - \pi_0 = \rho \tag{2.1}$$

– **Le nombre moyen de clients dans la file**

$$\begin{aligned} \overline{N}_Q &= \sum_{n \geq 1} (n-1)\pi_n \\ &= \frac{\rho^2}{1-\rho} \end{aligned}$$

Temps moyen de séjour $\overline{T}$: Le temps moyen de séjour $\overline{T}$ se calcule en utilisant la loi de Little :

$$\begin{aligned} \overline{T} &= \overline{N}/\lambda \\ &= \frac{\rho}{\lambda(1-\rho)} \\ &= \frac{1}{\mu-\lambda} \end{aligned}$$

– **Temps moyen de service**

$$\overline{T}_S = 1/\mu \qquad (2.2)$$

– **Temps moyen d'attente**

$$\begin{aligned} \overline{T}_Q &= \overline{T} - \overline{T}_S \\ &= \frac{\lambda}{\mu(\mu-\lambda)} \end{aligned}$$

2.7 File $M/M/1/K$

Soit K la capacité de la file d'attente : c'est le nombre maximal de clients qui peuvent être présents dans le système, soit en attente, soit en service. Quand un client arrive alors qu'il y a déjà K clients présents dans le système, il est perdu. Ce système est connu sous le nom de file $M/M/1/K$. L'espace d'états E est maintenant infini : $E = \{0, 1, 2, \ldots\}$ La capacité de la file étant limitée, même si les clients arrivent en moyenne beaucoup plus vite que ce que le serveur de la file est capable de traiter, dès que celle-ci est pleine, les clients qui se présentent sont rejetés. Le nombre de clients dans la file ne peut donc jamais tend à l'infini. De plus, dès qu'un client est autorisé à entrer, il sortira un jour et son temps de séjour dans la file est fini, puisqu'il correspond au temps de service de tous les clients devant lui et que ce nombre est limité par K. Sur un temps très long, le débit de sortie sera donc bien égal au débit d'entrèe, ce qui correspond bien à la stabilité inconditionnelle du système.

On défini Le processus de naissance et de mort qui modéliser ce type de file d'attente par :

$$\lambda_n = \begin{cases} \lambda, & \text{si } n < K\,; \\ 0, & \text{si } n > K\,; \end{cases}$$

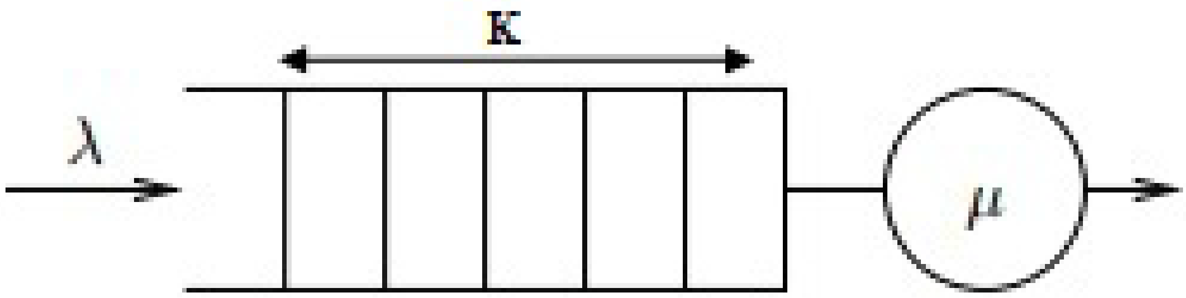

FIGURE 2.6 – La file $M/M/1/K$

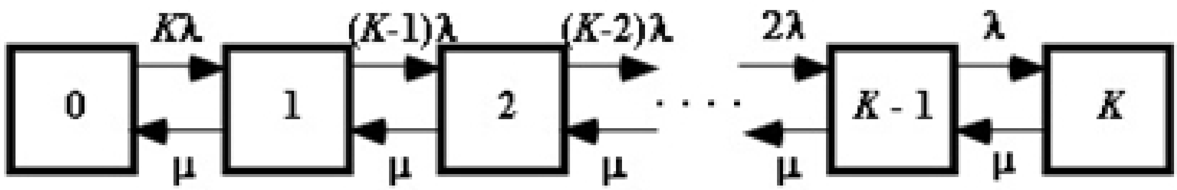

FIGURE 2.7 – Evaluation de l'état dans la file d'attente $M/M/1/K$

L'intégration de l'équation récurrente permettant de calculer π_n se fait alors comme suit :

$$\begin{aligned} \pi_n &= \pi_0 \rho^n \quad \text{pour} \quad n \leq K \\ \pi_n &= 0 \quad \text{pour} \quad n > K \end{aligned}$$

et

$$\pi_0 = \begin{cases} \dfrac{1}{\sum_{n=0}^{K} \rho^n} = \dfrac{1-\rho}{1-\rho^{K+1}}, & \text{si} \quad \lambda \neq \mu; \\ \dfrac{1}{K+1} & \text{si} \quad \lambda = \mu. \end{cases}$$

– Le nombre moyen de clients dans le système est :

$$\begin{aligned} \overline{N} &= \sum_{n=0}^{K} n\pi_n \\ &= \frac{\rho}{1-\rho} \frac{1-(K+1)\rho^K + K\rho^{K+1}}{1-\rho^{K+1}} \end{aligned}$$

Lorsque K tend vers l'infini et $\rho < 1$, c'est le cas de la file $M/M/1$:

$$\overline{N} = \frac{\rho}{1-\rho}$$

– Le nombre moyen de clients dans la file est :

$$\begin{aligned} \overline{N}_Q &= \sum_{n=1}^{\infty}(n-1)\pi_n \\ &= \overline{N} - (1-\pi_0) \end{aligned}$$

à partir de la loi de Little on obtien le temps moyen qu'un client passe dans le système $\overline{T}$ et le temps moyen d'attente dans la file $\overline{T}_Q$

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

$$\overline{T} = \frac{\overline{N}}{\lambda} \tag{2.3}$$

– Temps moyen d'attente

$$\overline{T}_Q = \frac{\overline{N}_Q}{\lambda} \tag{2.4}$$

2.8 File $M/M/C$

On considère un système identique à la file $M/M/1$ excepté qu'il comporte C serveurs identiques et indépendants les uns des autres. On conserve les hypothèses :

– Le processus d'arrivée des clients poissonien de taux λ

– Le temps de service exponentiel de taux μ

Ce système est connu sous le nom de la file $M/M/C$. L'espace d'états E est, comme pour la $M/M/1$ infini : E (0, 1,2, ...) La file d'attente est de capacité infini. Si l'un des serveurs est libre, le client qui arrive se dirige immédiatement vers ce serveur. Dans le cas contraire, le client prend sa place dans une file d'attente commune pour tous les serveurs. Lorsqu'un serveur se libère, le client en tête de la file occupe ce serveur. Par conséquent, la discipline d'attente est FIFO

Le processus de naissance et de mort modélisant ce type de file d'attente est alors défini de la façon suivante :

$$\lambda_n = \lambda$$

$$\mu_n = \begin{cases} 0n\mu & n = 1, \ldots, C-1 \\ C\mu & n \geq C \end{cases}$$

La condition de stabilité de ce modèle est :$\rho = \dfrac{\lambda}{C\mu} < 1$

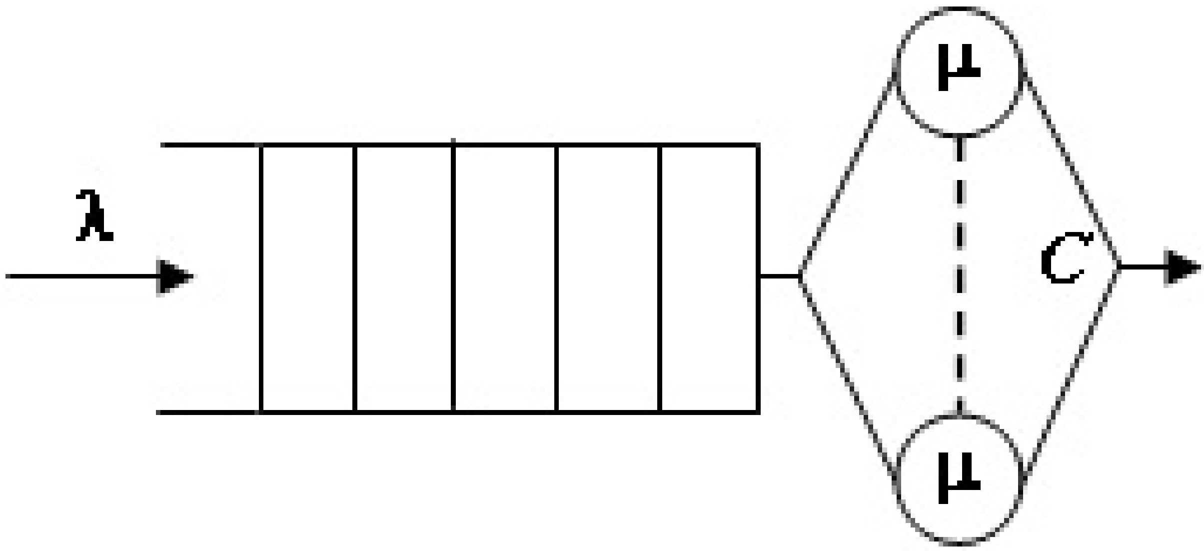

FIGURE 2.8 – La file $M/M/C$

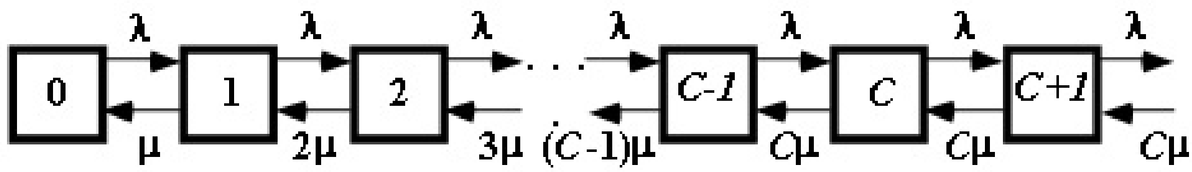

FIGURE 2.9 – Evaluation de l'état dans la file d'attente $M/M/C$.

D'aprè le diagramme et l'analyse du système en régime stationnaire, à l'aide de la procédure des équations de Chapman Kolmogorov on obtien les equation équations suivantes :

$$\lambda\pi_0 = \mu\pi_1$$
$$(\lambda + n\mu)\pi_n = \lambda\pi_{n-1} + (n+1)\mu\pi_{n+1} \quad 1 \leq n < c$$
$$(\lambda + c\mu)\pi_n = \lambda\pi_{n-1} + c\mu\pi_{n+1} \quad n \geq c$$

avec

$$\sum_{n=0}^{\infty} \pi_n = 1$$

La résolution du système ci-dessus présente la distribution stationnaire suivante :

$$\begin{aligned} \overline{N}_Q &= \sum_{n=1}^{\infty} (n-1)\pi_n \\ &= \overline{N} - (1 - \pi_0) \end{aligned}$$

$$\pi_n = \frac{\rho^C}{C!}(A)^{n-C}\pi_0, \quad n \geq C \tag{2.5}$$

où

$$\pi_0 = \left[\sum_{n=0}^{C-1} \frac{\rho^n}{n!} + \frac{\rho^C}{C!}\sum_{n=C}^{\infty} \rho^{n-C}\right]^{-1}$$

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

et

$$A = \frac{\lambda}{C\mu}$$

Cette dernière existe si : $\lambda < C\mu$

A partir de la distribution stationnaire du processus $\{N(t), t \geq 0\}$, on peut calculer les caractéristiques du système. En effet,

– Le nombre moyen de clients dans le système :

$$\overline{N} = \rho + \frac{\rho^{C+1}}{C.C!(1-A)^2}\rho_0 \tag{2.6}$$

– Le nombre moyen de clients dans la file :

$$\overline{N}_Q = \frac{\rho^{C+1}}{C.C!(1-A)^2}\rho_0 \tag{2.7}$$

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

$$\overline{T} = \frac{C\mu\rho^C}{C!(C\mu-\lambda)^2}\rho_0 \tag{2.8}$$

– Temps moyen d'attente :

$$\overline{T}_Q = \frac{1}{\mu} + \frac{\rho^C}{\mu C \cdot C!(1-A)^2}\rho_0 \tag{2.9}$$

2.9 File $M/M/\infty$

Pour ce modèle de file d'attente, le système est composé d'un nombre illimité de serveurs identiques et indépendants les uns des autres. Dés qu'un client arrive, il est immédiatement servi (c'est le cas ou il n'y a pas d'attente). Dans cette file les clients arrivent à des instants$0 < t1 < t2 <$ ...formant un processus de Poisson de taux λ et les temps de service sont exponentiels de taux μ Ce système est connu sous le nom de file $M/M/\infty$ Comme cela a été fait pour la file $M/M/C$, on peut facilement démontrer que le taux de transition d'un état n quelconque vers l'état $n-1$ est égal à $n\mu$ et correspond au taux de sortie d'un des n clients en service. De même, le taux de transition d'un état n vers l'état $n+1$ est égal à λ et correspond au taux d'arrivée d'un client. donc c'est un processus de naissance est de mort avec :

$$\pi_{n-1}\lambda = \pi_n n\mu \quad \text{pour} \quad n = 1, 2, \ldots$$

$$\text{soit} \quad \pi_n = \frac{\rho}{n}\pi_{n-1} \quad \text{pour} \quad n = 1, 2, \ldots,$$

$$\text{où} \quad \rho = \frac{\lambda}{\mu}$$

les probabilités en fonction de π_n.

$$\pi_n = \frac{\rho^n}{n!}\pi_0 \quad \text{pour} \quad n = 1, 2, \ldots$$

La condition de normalisation nous donne alors immédiatement π_0

$$\pi_0 = \frac{1}{\displaystyle\sum_{n=0}^{+\infty} \frac{\rho^n}{n!}} = e^{-\rho}$$

On obtient finalement

$$\pi_n = \frac{\rho^n}{n!}e^{-\rho} \quad \text{pour} \quad n = 1, 2, \ldots$$

car la série $\displaystyle\sum_{n=0}^{+\infty} \frac{\rho^n}{n!}$ converge pour toutes valeurs de ρ (donc de λ et de μ), ce qui est cohérent avec la stabilité inconditionnelle de la file

– Nombre moyen de clients $\overline{N}$:

$$\begin{aligned}\overline{N} &= \sum_{n=1}^{+\infty} n\pi_n \\ &= e^{-\rho}\sum_{n=1}^{+\infty}\frac{\rho^n}{(n-1)!} \\ &= e^{-\rho}\rho e^{\rho} = \rho\end{aligned}$$

– Temps moyen de séjour $\overline{T}$:
utilisant la loi de Little :

$$\begin{aligned}\overline{T} &= \frac{\overline{N}}{\lambda 0} \\ &= \frac{1}{\mu}\end{aligned}$$

Chapitre 3

Système de files d'attente avec vacance simple de serveur

les serveurs peuvent ne plus être disponibles pendant un certain temps pour diverses raisons. Cette période d'absence du serveur peut signifier que le serveur travaille sur certaines tâches supplémentaires, pour analyser ces systèmes, nous introduisons les vacances de service dans les modèles de files d'attente afin de représenter la période d'absence temporaire du serveur et ça rend les modèles de file d'attente plus réalistes et plus flexibles. Depuis que l'idée a été discutée pour la première fois dans l'article de Levy et Yechiali [12] .Les systèmes de file d'attente avec des vacances sur serveur ont attiré l'attention de nombreux chercheurs. Plusieurs enquêtes sur ces modèles de vacances ont été réalisées par Doshi [4], [3] et les livres de Takagi [7], Tian et Zhang [11] sont consacrés à ce sujet.

3.1 Vacances dans le système de files d'attente

Un système de files d'attente avec vacances est un système dans lequel un serveur peut devenir indisponible pendant une période aléatoire à partir d'un centre de service principal. On appelle Le temps passé loin du centre de service principal les vacances et peut être le résultat de nombreux facteurs. Dans certains cas, les vacances peuvent résulter d'une panne du serveur, ce qui signifie que le système doit être réparé et remis en service. Il peut également s'agir d'une action de libéré le serveur pour l'utiliser dans un centre de service secondaire lorsqu'il n'y a aucun client présent dans le centre de service principal. Ainsi, les vacances de serveur sont utiles pour les systèmes dans lesquels le serveur souhaite utiliser son temps d'inactivité à différentes fins, ce qui permet d'appliquer le modèle de files d'attente à une variété de systèmes de service stochastiques du monde réel.

3.2 Système de files d'attente avec vacance simple

Le serveur prend exactement une vacance dès que le système est vide. S'il trouve un système vide à son retour de vacance, il devient inactif jusqu'à l'arrivée d'un client. Par exemple, la maintenance des machines dans un processus de production est considérée comme une vacance simple.

3.2.1 Description mathématique du modèle

Nous considérons un système de file d'attente dans lequel les clients arrivent selon un processus de Poisson au taux λ. Le service est fourni par un serveur unique de capacité illimitée. Le temps de service successif de le serveur est distribué de manière exponentielle avec un débit μ. Chaque fois que le système devient vide, le serveur prend des vacances. Si au retour de vacances, le système est vide, le serveur attend pour la prochaine arrivée, sinon il commence le service. Les temps de vacances sont répartis de manière exponentielle au taux γ. Supposons en outre que les clients sont impatients. Chaque fois qu'un client arrive au système, il s'impatiente quels que soient les états du serveur. Les temps impatients sont distribué de façon exponentielle au taux ξ. Si un client rejoint la file d'attente, il reste jusqu'à ce que son service soit terminé.

Le système est représenté par une chaîne de Markov continue bidimensionnelle $J(t), N(t)$ avec espace d'états $S = \{(j,n) : j = 0,1 \quad n = 0,1,2.....\}$ où $N(t)$ désigne le nombre total de clients présent dans le système et $J(t)$ indique l'état du système au temps t. Si $J(t) = 0$, le serveur est en vacances, si $J(t) = 1$, le serveur sert des clients ou il est inactif. Le diagramme de transition est représenté sur la figure (3.2.1).
Définissons les probabilités stationnaires pour la chaîne de Markov comme :

$$p_{j,n} = P(J(t) = j, N(t) = n), j = 0,1, n = 0.$$

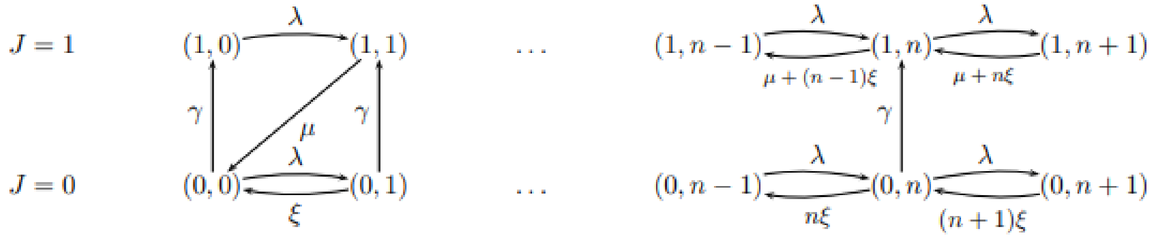

FIGURE 3.1 – Diagramme du taux de transition de $(J(t), N(t))$ pandent des vacances uniques

3.2.2 Distribution de l'état d'équilibre

À l'aide du procédé des équations de Chapman-Kolmogorov aboutissent aux équations d'équilibre suivantes :

$$(\lambda + \gamma)p_{0,0} = \xi p_{0,1} \mu p_{1,1} \tag{3.1}$$

$$\lambda + \gamma + n\xi p_{0,n} = \lambda p_{0,n-1} + (n+1)\xi p_{0,n+1} \tag{3.2}$$

$$\lambda p_{1,0} = \gamma p_{0,0} \tag{3.3}$$

$$(\lambda + \mu + (n-1)\xi)p_{1,n} = \lambda p_{1,n-1} + \gamma p_{0,n} + (\mu + n\xi)p_{1,n+1} \tag{3.4}$$

Il est pratique d'introduire la fonction génératrice de probabilité (PGF) afin de résoudre les équations (3.1) - (3.4). Définissez les fonctions génératrices de probabilité (partielle) (PGF) P_0 (z) et$P_1(z)$, pour $0 < z < 1$

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

avec

$$P_0(1) + P_1(1) = 1.$$

Notons $P_0'(z)$ et$P_1'(z)$ les dérivées respectives de $P_0(z)$ et$P_1(z)$, pour $0 < z < 1$

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

Les cas particuliers :

1. $\xi = 0$ Ce cas correspond à un système où les clients ne sont pas impatients. Ce modèle se résume à un $M/M/1/Vs$.
2. $\xi = 0$ uniquement dans l'équation (3.4). Ce système correspond à un système où les clients ne sont impatients que lorsque le serveur est en vacances. Cette affaire a été résolue dans un premier temps par Altman et Yechiali [12] [2006] pour les files d'attente $M/M/1$, $M/G/1$ et $M/M/c$ avec des vacances multiples et uniques
3. $\gamma = 0$ Ce cas représente un système sans vacances. L'espace d'états est alors unidimensionnel et le système se résume à une file d'attente $M/M/1 + M$. Un certain nombre d'articles sur les files d'attente avec des phénomènes d'impatience sont apparus. On peut citer par exemple, pour la file d'attente $M/M/c$ et $M/M/1$

Ensuite, on va trouver les fonctions génératrices P_0 et P_1. En multipliant (3.1)(3.2) par z^n et sommation sur n, on obtient, pour $0 < z < 1$

$$\xi(1-z)P_0^{'}(z) - [\lambda(1-z)+\gamma]P_0(z) = -\mu p_{0,0} \tag{3.5}$$

L'équation (3.5) est similaire à l'équation (5.5) dans Altman et Yechiali [2006][12] et a été résolue par le même auteur. On obtient de la même manière une équation différentielle du premier ordre liée à P_1, en utilisant (3.3)et (3.4), nous avons

$$(\lambda z - \mu + \xi)(1-z)P_1(Z) - \xi Z(1-Z)P_1^{'}z = \gamma z P_0(Z)(\xi-\mu)(1-z)p_{1,0} - \mu z p_{1,1} \tag{3.6}$$

Nous obtenons un ensemble d'équations différentielles linéaires du premier ordre impliquant les probabilités inconnues $p_{1,0}$ et $p_{1,1}$. Nous verrons en résolvant (3.5) que $p_{1,1}$ sera exprimé en termes de $p_{0,0}$. De plus, en utilisant l'équation (3.3), $p_{1,0}$ est exprimé en termes de $p_{0,0}$. Ainsi, les PGF P_0 et P_1 seront exprimées en termes de $p_{0,0}$. Par conséquent, une fois que $p_{0,0}$ sera calculé, P_0 et P_1 seront complètement déterminés.

Diviser (3.6) par $\xi(1-z)$ donne

$$P_0^{'}(z) - \left[\frac{\lambda}{\xi} + \frac{\gamma}{\xi(1-z)}\right]P_0(z) = -\frac{\mu}{\xi(1-z)}p_{1,1} \tag{3.7}$$

Pour résoudre l'équation différentielle du premier ordre ci-dessus, un facteur d'intégration peut être trouvé comme

$$exp\left(\int -(\frac{\lambda}{\xi} + \frac{\gamma}{\xi(1-z)})dz\right) = exp\left(-\frac{\lambda}{\xi}z\right)(1-z)^{\frac{\gamma}{\xi}} \tag{3.8}$$

Multiplier les deux côtés de (3.7) par le facteur d'intégration $e^{-\frac{\lambda}{\xi}z}(1-z)^{\frac{\gamma}{\xi}}$

$$\frac{d}{dz}\left[exp\left(-\frac{\lambda}{\xi}z\right)(1-z)^{\frac{\gamma}{\xi}}P_0(z)\right] = -p_{(1,1)}\frac{\mu}{\xi}exp\left(-\frac{\lambda}{\xi}z\right)(1-z)^{\frac{\gamma}{\xi}-1}$$

En intégrant de 0 à z, nous avons

$$exp\left(-\frac{\lambda}{\xi}z\right)(1-z)^{\frac{\gamma}{\xi}}P_0(z) - P_0(0) = -p_{(1,1)}\frac{\mu}{\xi}\int_0^z exp\left(-\frac{\lambda}{\xi}s\right)(1-s)^{\frac{\gamma}{\xi}-1}ds$$

Par conséquent

$$P_0(z) = exp\bigg(-\frac{\lambda}{\xi}z\bigg)(1-z)^{\frac{\gamma}{\xi}}\left[P_0(0) - p_{(1,1)}\frac{\mu}{\xi}\int_0^z exp\bigg(-\frac{\lambda}{\xi}s\bigg)(1-s)^{\frac{\gamma}{\xi}-1}ds\right] \tag{3.9}$$

Puisque $P_0(1) = \sum_{n=0}^{\infty} p_{0,n} < \infty$ et $\lim_{z\to 1}(1-z)^{\frac{\gamma}{\xi}} = \infty$

en prenant la limite z tendant à 1 dans (3.9) , et en notant que $P_0(0) = p_{0,0}$ on exprime $p_{1,1}$ en termes de $p_{0,0}$

$$p_{1,1} = \frac{\xi}{\mu A}p_{0,0} \tag{3.10}$$

avec

$$A(z) = \int_0^z exp\bigg(-\frac{\lambda}{\xi}s\bigg)(1-s)^{\frac{\gamma}{\xi}-1}ds \tag{3.11}$$

et

$$A = A(1) = \int_0^1 exp\bigg(-\frac{\lambda}{\xi}s\bigg)(1-s)^{\frac{\gamma}{\xi}-1}ds \tag{3.12}$$

Notez que $A(z)$ est bien défini si $\frac{\gamma}{\xi} - 1 > 0$ .La substitution de (3.10) (3.11)

$$P_0(z) = p_{0,0}exp\bigg(\frac{\lambda}{\xi}z\bigg)(1-z)^{-\frac{\gamma}{\xi}}\left[1 - \frac{A(z)}{A}\right] \tag{3.13}$$

qui est la même équation (2.12) [12], avec $p_{1,1}$ donné par (3.10). L'équation (3.13) exprime $P_0(z)$ en termes de $p_{0,0}$. Ainsi, une fois que $p_{0,0}$ est calculé, $P_0(z)$ est complètement déterminé. Le théorème suivant se concentre sur le calcul de $p_{0,0.}$ Ensuite, nous trouverons la fonction génératrice P_1. Diviser (3.6)par $-\xi z(1-z)$ nous donne

$$P_1'(z) - \left[\frac{\lambda}{\xi} - (\frac{\mu}{\xi} - 1)\frac{1}{z}\right] = -\frac{\gamma}{\xi(1-z)}P_0(z) - \frac{1}{z}(1-\frac{\mu}{\xi})p_{1,0} + \frac{\mu}{\xi(1-z)}p_{1,1}$$

En utilisant l'équation (3.3), ainsi que (3.10), l'équation ci-dessus s'écrit

$$P_1'(z) - \left[\frac{\lambda}{\xi} - (\frac{\mu}{\xi} - 1)\frac{1}{z}\right] = -\frac{\gamma}{\xi(1-z)}P_0(z) + \left[\frac{(\mu-\xi)\gamma}{\xi\lambda z} + \frac{1}{(1-z)A}\right]p_{0,0} \tag{3.14}$$

Multiplier les deux côtés de (3.14) par $e^{-\frac{\lambda}{\xi}z}z^{\frac{\mu}{\xi}-1}$ avec $\frac{\mu}{\xi}>0$ et integrer de 0 a z la soluion de l'equation déefferential est donnée par

$$p_1(z)=e^{-\frac{\lambda}{\xi}z}z^{\frac{\mu}{\xi}-1}\int_0^z e^{-\frac{\lambda}{\xi}z}z^{\frac{\mu}{\xi}-1}\left[-\frac{\gamma}{\xi(1-s)}P_0(s)+\left[\frac{(\mu-\xi)\gamma}{\xi\lambda s}+\frac{1}{(1-s)A}\right]p_{0,0}\right]$$

où $P_0(.)$ est donné par (3.13). L'équation ci-dessus est valide si et seulement si $\mu/\xi>1$ ce qui conduit à $\mu>\xi$ En remplaçant (3.13) dans l'équation ci-dessus, nous aurons

$$P_1(z)=p_{0,0}exp\left(-\frac{\lambda}{\xi}z\right)z^{\frac{\mu}{\xi}-1}\left[-\frac{\gamma}{\xi}B(z)+\frac{(\mu-\xi)\gamma}{\xi\lambda}C(z)+\frac{D(z)}{A}\right] \tag{3.15}$$

où

$$B(z)=\int_0^z s^{\frac{\mu}{\xi}-1}(1-s)^{-(\frac{\gamma}{\xi}+1)}\left(1-\frac{A(s)}{A}\right) \tag{3.16}$$

$$C(z)=\int_0^z e^{-\frac{\lambda}{\xi}s}s^{\frac{\mu}{\xi}-2}ds \tag{3.17}$$

et

$$D(z)=\int_0^z e^{-\frac{\lambda}{\xi}s}s^{\frac{\mu}{\xi}-1}(1-s)^{-1}ds \tag{3.18}$$

Tout ce qui précède est résumé dans la proposition suivante.

Proposition 3.2.2.1. *si $\rho=\frac{\lambda}{\mu}<1$ et $\frac{\xi}{\mu}$ <1 la fonction génératrice partielle peut être exprimée en termes de $p_{0,0}$ en tant que*

$$P_0(z)=p_{(0,0)}exp\left(\frac{\lambda}{\xi}z\right)(1-z)^{-\frac{\gamma}{\xi}}\left[1-\frac{A(z)}{A}\right] \tag{3.19}$$

$$P_1(z)=p_{(0,0)}exp\left(-\frac{\lambda}{\xi}z\right)z^{\frac{\mu}{\xi}-1}\left[-\frac{\gamma}{\xi}B(z)+\frac{(\mu-\xi)\gamma}{\xi\lambda}C(z)+\frac{D(z)}{A}\right] \tag{3.20}$$

Les équations (3.19) et (3.20) sont exprimées en termes de $p_{0,0}$, la probabilité qu'il n'y ait pas de client dans le système lorsque le serveur est en vacances. Afin de déterminer complètement $P_0(.)$ et $P_1(.)$, nous doivent calculer $p_{0,0}$. La proposition suivante donne cette valeur.

Proposition 3.2.2.2. *Si $\rho = \frac{\lambda}{\mu} < 1$ et $\frac{\xi}{\mu}$ <1 la probabilitè $p_{0,0}$ est donnée par*

$$p_{0,0}(z) = \left[\frac{\xi}{\gamma A} + e^{\frac{\lambda}{\xi}}\left\{-\frac{\gamma}{\xi}B(1) + \frac{(\mu-\xi)\gamma}{\xi\lambda}C(1) + \frac{D(1)}{A}\right\}\right]^{-1} \tag{3.21}$$

Preuve :

L'idée consiste à exprimer $P_0(1)$ et $P_1(1)$, qui sont respectivement la probabilité que le serveur est en vacances et la probabilité que le serveur serve des clients ou que le serveur soit idle, en termes de $p_{0,0}$. Pour obtenir $p_{0,0}$, nous utilisons la condition de normalisation De l'équation (3.19), en utilisant la règle de L'Hopital, on obtient

$$P_0(1) = \lim_{z\to 1} \frac{e^{\frac{\lambda}{\xi}} z p_{0,0}\left\{\frac{\lambda}{\xi}(1 - A^{-1}A(z)) - A^{-1}A'(z)\right\}}{-\frac{\gamma}{\xi}(1-z)^{\frac{\gamma}{\xi}} - 1}$$

où $A'(z) = e^{-\frac{\lambda}{\xi}z}(1-z)^{\frac{\gamma}{\xi}-1}$

$$P_0(1) = p_{0,0}\frac{\xi}{\gamma A} \tag{3.22}$$

utilisons (3.20) pour $z = 1$

$$P_1(1) = 1 - P_0(1) = p_{(0,0)} exp\left(\frac{\lambda}{\xi}\right)\left\{-\frac{\gamma}{\xi}B(1) + \frac{(\mu-\xi)\gamma}{\xi\lambda}C(1) + \frac{D(1)}{A}\right\}$$

utilisons la condition de normalisation $P_0(1) + P_1(1) = 1$, on aura

$$p_{0,0}(z) = \left[\frac{\xi}{\gamma A} + e^{\frac{\lambda}{\xi}}\left\{-\frac{\gamma}{\xi}B(1) + \frac{(\mu-\xi)\gamma}{\xi\lambda}C(1) + \frac{D(1)}{A}\right\}\right]^{-1}$$

3.2.3 Mesure de performance

La probabilité que le serveur soit en vacance.

$$\begin{aligned} P_{vac} &= \sum_{n=0}^{\infty} \mathbb{P}(J = 0, N = n) \\ &= \sum_{n=0}^{\infty} p_{0,n} \\ &= P_0(1) = p_{0,0}\frac{\xi}{\gamma A} \end{aligned}$$

avec $p_{0,0}$ donner par (3.21)

La probabilité que le serveur soit en service.

$$P_{idle} = \mathbb{P}(J = 1, N = 0) = p_{0,1} = \frac{\gamma}{\lambda} p_{0,0}$$

La probabilité que le serveur soit au service des clients.

$$\begin{aligned}
P_{ser} =& \sum_{n=1}^{\infty} \mathbb{P}(J = 1, N = n) \\
=& \sum_{n=0}^{\infty} p_{1,n} \\
=& 1 - P_{idel} - P_{vac} \\
=& 1 - p_{0,0} \left[\frac{\gamma}{\lambda} - \frac{\xi}{\gamma A(1)} \right]
\end{aligned}$$

Pour $j = 0, 1$, N_j la taille du système lorsque le serveur est dans l'état j. Ensuite, $\mathbb{E}(N_j)$ est la moyenne de la taille du système lorsque le serveur est dans l'état j

Le nombre moyen de clients lorsque le système est en vacances :

$$\mathbb{E} = (N_0) \sum_{n=0}^{\infty} n\mathbb{P}(J = 0, N = n) = \sum_{n=1}^{\infty} np_{0,n} = P_0^{'}(1)$$

De (3.7)et en utilisant la règle de l'hôpital, nous obtenons

$$\begin{aligned}
P_0^{'}(1) =& \lim_{z \to 1} \frac{[\lambda(1 - z) + \gamma] P_0(z) - \mu p_{1,1}}{\xi(1 - z)} \\
=& \lim_{z \to 1} \frac{-\lambda P_0(z) + [\lambda(1 - z) + \gamma] P_0^{'}(z)}{-\xi} \\
=& \frac{\lambda P_0(1) + \gamma P_0^{'}(1)}{-\xi}
\end{aligned}$$

Ainsi, nous obtenons,

$$P_0^{'}(1) = \frac{\lambda}{\gamma + \xi} P_0(1) = \frac{\lambda}{\gamma + \xi} \frac{\xi}{\gamma A} p_{0,0} \tag{3.23}$$

Nous déduisons maintenant le nombre moyen de clients pendant la période de non-vacance. De (3.6), nous obtenons

$$\mathbb{E}(N_1) = P_1^{'}(z) = \frac{(\lambda z - \mu + \xi)(1-z) - \gamma z P_0(z) - (\xi - \mu)(1-z)p_{1,0} + \mu z p_{1,1}}{\xi z(1-z)}$$

en utilisant la règle de l'hopital nous onbtenons

$$P_1^{'}(1) = \lim_{z \to 1} \frac{\lambda(1-z) - (\lambda z + \mu + \xi) - \gamma z P_0(z) + \gamma z P_0^{'}(z) - (\xi - \mu)p_{1,0} + \mu p_{1,1}}{\xi(1-2z)}$$

où $p_{0,0}$ est donné par (3.21).

$$\mathbb{E} = (N_0) \sum_{n=0}^{\infty} n\mathbb{P}(J = 0, N = n) = \sum_{n=1}^{\infty} n p_{0,n} = P_0^{'}(1)$$

De (3.7)et en utilisant la règle de l'hôpital, nous obtenons

$$\begin{aligned} P_0^{'}(1) &= \lim_{z \to 1} \frac{\left[\lambda(1-z) + \gamma\right] P_0(z) - \mu p_{1,1}}{\xi(1-z)} \\ &= \lim_{z \to 1} \frac{-\lambda P_0(z) + \left[\lambda(1-z) + \gamma\right] P_0^{'}(z)}{-\xi} \\ &= \frac{\lambda P_0(1) + \gamma P_0^{'}(1)}{-\xi} \end{aligned}$$

Ainsi, nous obtenons,

$$P_0^{'}(1) = \frac{\lambda}{\gamma + \xi} P_0(1) = \frac{\lambda}{\gamma + \xi} \frac{\xi}{\gamma A} p_{0,0} \tag{3.24}$$

Nous déduisons maintenant le nombre moyen de clients pendant la période de non-vacance. De (3.6), nous obtenons

$$\mathbb{E}(N_1) = P_1^{'}(z) = \frac{(\lambda z - \mu + \xi)(1-z) - \gamma z P_0(z) - (\xi - \mu)(1-z)p_{1,0} + \mu z p_{1,1}}{\xi z(1-z)}$$

en utilisant la règle de l'hopital nous onbtenons

$$\begin{aligned}
P_1'(1) =& \lim_{z\to 1} \frac{\lambda(1-z) - (\lambda z + \mu + \xi) - \gamma z P_0(z) + \gamma z P_0'(z) - (\xi - \mu)p_{1,0} + \mu p_{1,1}}{\xi(1-2z)} \\
=& \frac{-(\lambda - \mu - \xi) - \gamma P_0(1) + \gamma P_0'(1) + (\xi - \mu)p_{1,0} + \mu p_{1,1}}{-\xi} \\
=& \left[-(\lambda z + \mu + \xi) - \frac{\lambda}{\gamma + \xi}\frac{\xi}{\gamma A} p_{0,0} \frac{\xi\gamma}{\lambda} p_{0,0} - \frac{\mu\gamma}{\lambda} p_{0,0} \right] \times \frac{1}{-\xi}
\end{aligned}$$

$$\mathbb{E}(N_1) = \frac{\lambda - \mu + \xi}{\xi} + \left[\frac{\lambda}{(\lambda + \xi)} - \frac{\gamma}{\lambda} + \frac{\mu\gamma}{\lambda\xi} \right] p_{0,0} \tag{3.25}$$

où $p_{0,0}$ est donné par (3.21).

Temps de séjour

Soit S le temps de séjour total d'un client arbitraire dans le système, indépendamment qu'il a terminé ou non son service. Soit $S_{j,n} = \mathbb{E}(S|(X_0 = (j, n+1)))$ le temps de séjour conditionnel d'un client qui n'abandonne pas le système, étant donné que l'état à son arrivée est (j, n), car le client marqué est inclus dans le système. Nous utilisons les chaînes de Markov. La durée totale du séjour lorsque le client est marqué à son arrivée est $(1, 0)$, c'est-à-dire que le serveur est idle, est donnée par

$$\mathbb{E}(S_{1,0}) = \frac{1}{\mu}$$

Calculons $\mathbb{E}(S_{1,n})$. En conditionnant la prochaine transition à un départ (soit en raison d'un service terminé ou d'un client impatient) ou une arrivée, nous obtenons

$$\begin{aligned}
\mathbb{E}(S_{1,n}) =& \mathbb{E}(S|(X_0 = (1, n+1))) \\
=& \frac{\mu + n\xi}{\beta_n} \mathbb{E}(S|X_0 = (1, n+1), X_1 = (1, n)) + \frac{\lambda}{\beta_n} \mathbb{E}(S|X_0 = (1, n+1), X_1 = (1, n+2)) \\
=& \frac{\mu + n\xi}{\beta_n} \left(\frac{1}{\beta_n} + \mathbb{E}(S|X_0 = (1, n)) \right) + \frac{\lambda}{\beta_n} \left(\frac{1}{\beta_n} + \mathbb{E}(S|X_0 = (1, n+1)) \right) \\
=& \frac{1}{\beta_n} + \frac{\mu + n\xi}{\beta_n} \mathbb{E}(S|X_0 = (1, n+1) + \frac{\lambda}{\beta_n} \mathbb{E}(S_{1,n}))
\end{aligned}$$

où $\beta_n = \lambda + \mu + n\xi$. Avec

$$\begin{aligned}\frac{\mu+n\xi}{\beta_n}\mathbb{E}(S|X_0=(1,n)) =& \frac{\mu}{\beta_n}\mathbb{E}(S_{1,n-1}) + \frac{n\xi}{\beta_n}\mathbb{E}(S_{1,n-1}) \\ =& \frac{\mu}{\beta_n}\mathbb{E}(S_{1,n-1}) + \frac{n\xi}{\beta_n}\left(\frac{1}{n}\times 0 + \frac{n-1}{n}\mathbb{E}(S_{1,n-1})\right)\end{aligned}$$

nous obtenons

$$\mathbb{E}(S_{1,n}) = \frac{1}{\beta_n} + \frac{\mu+(n-1)\xi}{\beta_n}\mathbb{E}(S_{1,n-1}) + \frac{\lambda}{\beta_n}\mathbb{E}(S_{1,n})$$

donc

$$\mathbb{E}(S_{1,n}) = \frac{1}{\mu+n\xi} + \frac{\mu+(n-1)\xi}{\beta_n}\mathbb{E}(S_{1,n})$$

$$\mathbb{E}(S_{1,n}) = \frac{n+1}{\mu+n\xi}$$

Conclusion

Ce mémoire cible des systèmes de files d'attente avec vacance simple. Les multiples applications de ces systèmes les rendent des sujets de recherches très solicités. Dans ce mémoire nous nous sommes intéresses au modèle $M/M/1$ avec vacance simple. les modèles de système de files d'attente avec vacance simple ont une très grandes importance dans les domaines de télécommunications, centrales téléphoniques .... Pour ces modèles le serveur peut ne pas être disponible pendant une période de temps. Dans cette période d'absence en communication par exemple, le serveur peut être utilisé à réaliser d'autre taches.

Bibliographie

[1] Abdel-Karim Aboul-Hassan, Sherif I. Rabia and Ahmed Kadry, Analytical study of a discrete time retrial queue with balking customers and early arrival scheme, Alexandria Engineering Journal, 44 (2005), No. 6, 911-917

[2] Anisimov, V. V., Zakusilo, O. K., and Donchenko, V. S. 1987. Elements of Queueing Theory and Asymptotic System Analysis. Vishcha Shkola, Kiev (in Russian).

[3] B. Doshi, "Single server queues with vacations," in Stochastic Analysis of Computer and Communication Systems, H. Takag, Ed., pp. 217-265, Elsevier, 1990.

[4] B. T. Doshi, "Queueing systems with vacations-a survey," Queueing Systems : Theory and Applications, vol. 1, no. 1, pp. 29-66, 1986.

[5] Claudie Chabriac. Processus stochastiques et modélisation. (2012-2013).

[6] Claudie Hasseforder CHABRIAC , Eléments de Théorie des files d'attente, page05, Janvier 2008.

[7] H. Takagi, Queueing Analysis : A Foundation of Performance Analysis, vol. 1 of Vacation and Priority Systems, part 1, Elsevier Science Publishers B.V., Amsterdam, The Netherlands, 1991.

[8] K.B, GK, Techniques de modélisation : Méthodes analytiques.

[9] Khintchine, A. Y 1969. Mathematical Methods in the Theory of Queueing. Second edition, Hafner Publishing Company, New York (First edition : Griffin, London, 1960 ; Russian original : 1955). 859-866.

[10] Newell, G. F. 1982. Applications of Queueing Theory. Second edition, Chapman and Hall, London (First edition : 1971).

[11] N. Tian and Z. G. Zhang, Vacation Queueing Models : Theory and Applications, Springer, New York, NY, USA, 2006

[12] Y. Levy and U. Yechiali, "Utilization of idle time in an $M/G/1$ queueing system," Management Science, vol. 22, no. 2, pp. 202- 211, 1975.

[13] Zakhar Kabluchko, Stochastic Processes (Stochastik II), University of Ulm Institute of Stochastics, (2013-2014).