

Table des matières

1	Les processus stochastiques	9
1.1	Introduction	9
1.2	Processus de comptage	10
1.3	Processus de renouvellement	10
1.3.1	Processus à accroissements indépendants	11
1.4	Processus de Poisson	11
1.4.1	Loi de Poisson et loi exponentielle	12
1.5	Processus de naissance et de mort	14
1.5.1	Processus de naissance	15
1.5.2	Processus de mort	16
2	Système de files d'attente	17
2.1	Définitions	17
2.2	Description d'une file simple	18
2.2.1	Processus d'arrivé :	19
2.2.2	Temps de service :	19
2.3	Structure et discipline de la file	20
2.3.1	Nombre de serveurs :	20
2.3.2	Capacité de la file :	20
2.3.3	Discipline de service :	21
2.4	Notation de Kendall	21
2.5	Loi de Little	22
2.6	Mesure de performance d'une file d'attente	23
2.7	Arrivée avant un départ et départ avant une arrivée	23
2.8	Analyse en régime stationnaire	24
2.9	La file $M/M/1$	24
2.10	La file $M/M/1/K$	27
2.11	La file $M/M/C$	29
2.12	La file $M/M/\infty$	31

3	Système de files d'attente avec vacances de serveur	34
3.1	Système de files d'attente avec vacance unique et vacances multiples . . .	34
3.1.1	Système de files d'attente avec vacance unique	35
3.1.2	Système de files d'attente avec vacances multiples	35
3.2	Système de files d'attente avec deux vacances différentes	36
3.2.1	Description mathématique du modèle	36
3.2.2	Analyse stationnaire du modèle	36
3.2.3	Mesures de performance	41
3.2.4	Application numérique	43
	Bibliographie	49

Remerciement

En premier lieu, je tiens à remercier mon **Dieu** qui m'a donné le courage et la volonté pour réaliser ce travail.

Je tiens d'exprimer toute ma reconnaissance à mon directeur de mémoire et mon professeur Dr. **Kadi Mokhtar**, de m'avoir encadré orienté aidé et conseillé, et ainsi pour la confiance et la liberté qu'il m'a accordé, sans oublier ses efforts fournis pendant cette période.

Je le remercie profondément pour sa disponibilité, ses orientations et ses remarques fructueuses. Qu'il trouve ici l'expression de ma profonde gratitude.

Mes vifs remerciements vont aux *Mr* **A. Guendouzi**, pour l'aide considérable qui m'a donné à réaliser la partie simulation.

Je remercie *Dr.* **S. Idrissi**, d'avoir accepté à présider le jury de mon mémoire.

Je souhaite également remercier Dr **F. Benziadi**. et *Mr.* **M. Laouni**. de m'avoir fait l'honneur d'accepter d'être examinateurs de mon mémoire. Leurs remarques et précieux conseils ont abouti à une amélioration de ce mémoire.

J'adresse mes sincères remerciements à tous Les membres du laboratoire des Modèles Stochastiques, Statistique et Applications et à la tête le directeur *Pr.* **A. Kandouci**.

Je remercie également tous les enseignants de départements de Mathématiques.

Je remercie mes très chers parents, qui ont toujours été là pour moi et surtout mon **père** et ma **mère**.

Enfin je remercie mes collègues étudiants et étudiantes de Mathématiques pour les bons moments passés ensemble et aussi pour leur soutiens et leur encouragements. Et pour finir, merci à toutes les personnes que j'ai oublié de citer et qui m'ont permis de mener à bien ce mémoire.

<Merci à vous tous>

Dédicace

Rien n'est plus dur pour tout étudiant que de trouver là où je suis.

Je dédie ce modeste travail à mes chers parents qui m'ont donnée la possibilité de poursuivre mes études, pour leurs guides affectueux, pour l'espoir qu'ils me donnent, pour les conseils dans la vie et leurs soutiens durant mes années d'études avec patience, et j'espère que je puisse leur rendre le maximum de bonheur qu'ils m'ont offert et que dieu les bénisse. A mes sœurs. A mes frères. A mes cousins et cousines.

A tous mes amis ainsi qu'à tous mes camarades.

A tous mes chers professeurs durant toutes mes années d'études.

Enfin, à tous ceux qui me sont chères, qu'ils trouvent ici l'expression de mon profond respect et ma gratitude. À tous ceux qui m'aiment. À tous ceux que j'aime.

Je dédie ce mémoire.

Introduction générale

La théorie des files d'attente datent du début du XXème siècle par les travaux de l'ingénieur danois Agner Krarup Erlang (1878, 1929). Ses études sur le trafic téléphonique de Copenhague pour le mieux gérer afin de déterminer le nombre de circuits nécessaires pour fournir un service téléphonique acceptable, sont considérées comme la première brique dans cette théorie [4]. Cette théorie fournit un outil très puissant et efficace pour la modélisation des systèmes admettant un phénomène d'attente. Ensuite, les applications des files d'attente dans la modélisation des systèmes informatiques et aux réseaux de communication ont permis une évolution de cette théorie surtout dans l'évaluation des paramètres de performances des systèmes informatiques et aux réseaux de communication. Actuellement ce sont les applications dans le domaine de l'analyse de performance des réseaux (téléphone mobile, internet, multimédia, ...) qui suscitent le plus de travaux. Pour évaluer le rendement d'un système de files d'attente et améliorer son fonctionnement (minimiser le temps d'attente et le temps d'inactivité de l'installation), il faut calculer ces paramètres de performance. Cela nous sert à prendre des décisions par exemple sur le nombre de serveurs adéquat pour gérer le flux de demandes dans le système ou encore d'appréhender les effets d'une modification des conditions de fonctionnement.

Depuis les travaux d'Erlang [1] Un grand nombre d'applications dans tous les domaines ont été mis en œuvre et publiées. En 1953, David G. Kendall [1] a introduit la notation de Kendall pour décrire les caractéristiques d'un système de file d'attente. En 1957 d'une manière particulièrement élégante et efficace Jackson a traité certains réseaux de files d'attente. En 1961, Thomas L. Saaty [2], auteur de l'un des premiers livres complets sur la théorie des files d'attente. Ensuite c'est les contributions des mathématiciens Khintchine, Palm, Pollaczek et Kolmogorov qui ont vraiment poussés la théorie des files d'attente.

Dans ce mémoire on s'intéresse aux systèmes de files d'attente avec vacances de serveur. Ces vacances sont dues à plusieurs facteurs. Elles 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 délibérée visant à utiliser le serveur dans un centre de service secondaire lorsqu'il n'y a aucun client présent dans le centre de service principal.

Mon mémoire est composé de trois chapitres : Dans le premier chapitre, nous présen-

tons les notions de bases de l'étude des systèmes de files d'attente, à savoir les processus stochastiques :

- Processus de comptage,
- processus de renouvellement,
- processus de Poisson,
- processus de naissance et de mort.

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). Ensuite nous étudions quelque modèles de files d'attente ($M/M/1$, $M/M/1/K$, $M/M/C$, $M/M/\infty$) et l'évaluation de leurs paramètres de performance. Enfin dans le troisième chapitre nous présentons une étude d'un modèle de file d'attente avec vacances de serveur. Nous traitons le cas de files d'attente $M/M/1$ avec vacances multiples. Une étude numérique est donnée pour montrer l'influence de certain paramètres les uns sur les autres.

Chapitre 1

Les processus stochastiques

1.1 Introduction

Les processus stochastiques décrivent l'évolution d'une grandeur aléatoire en fonction du temps (ou de l'espace), il existe de nombreuses applications des processus aléatoires notamment en physique statistique [11]. Les processus stochastiques sont d'une importance majeure dans la modélisation des files d'attente, parmi ces processus :

le processus de Poisson, qui est un outil extrêmement utilisé dans les phénomènes de comptage. Il apparaît en effet naturellement dans ces situations, le deuxième est celui de **naissances et morts** qui sert à modéliser l'évolution d'une population au cours du 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 $\mathbf{X}(t)$ représente l'état du processus stochastique au temps t [18].

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

Remarque 1.1.1. *Dans les deux cas de la définition précédente, l'ensemble \mathbb{T} est muni d'une relation d'ordre \leq , qui de plus est 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 $\mathbb{T} = \{0, \dots, N\}$, 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*)

Un processus stochastique $N(t)$ est appelé processus de comptage si $N(t)$ représente le nombre total des événements qui sont arrivés avant l'instant t . Tout processus de comptage vérifie les propriétés suivantes :

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

Un processus de comptage est un processus discret à temps continu. Un second processus peut être associé au processus des temps d'occurrence; le processus des temps d'inter-arrivées $\{W_n, n \in \mathbb{N}_0\}$ ou $\forall n \in \mathbb{N}_0$ la variable aléatoire W_n est le temps d'attente entre les $(n-1)$ ième, n ième occurrences [17], c-à-d :

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

avec T_n est le temps d'arrivé du n ième client.

Propriété 1.2.1. Les relations suivantes sont triviales tel que $T_0 = 0$:

1. $T_n = W_1 + W_2 + \dots + W_n \quad \forall n \geq 1$;
2. $N(t) = \sup\{n \geq 0 : T_n \leq t\}$;
3. $\mathbb{P}[N(t) = n] = \mathbb{P}[T_n \leq t < T_{n+1}]$;
4. $\mathbb{P}[N(t) \geq n] = \mathbb{P}[T_n \leq t]$;
5. $\mathbb{P}[s < T_n < t] = \mathbb{P}[n(s) < n \leq N(t)]$.

Preuve on a

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

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

1.3 Processus de renouvellement

Un processus de renouvellement a pour fonction le dénombrement des occurrences d'un phénomène donné, lorsque les délais entre deux occurrences consécutives sont des variables aléatoires indépendantes et identiquement distribuées.

Exemple 1.3.1. *Il peut s'agir de compter le nombre de pannes d'un matériel électronique en théorie de la fiabilité (le matériel est alors renouvelé après chaque panne, d'où la dénomination), de dénombrer les arrivées de clients dans une file d'attente, de recenser les occurrences d'un sinistre pour une compagnie d'assurance...*

Définition 1.3.1. (Processus de renouvellement)

Un processus de comptage pour lequel les temps entre deux arrivées consécutives sont des variables aléatoires i.i.d. s'appelle processus de renouvellement. Les temps de renouvellement (ou les temps de la n -ième arrivée) sont :

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

avec $a_i, i = 1, 2, \dots$ est le temps entre deux arrivées consécutives. Il est facile de voir 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 1.3.1.1. (Processus à accroissements indépendants)

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

$$- \forall t \geq 0, \forall s \geq 0, N_{t+s} - N_t \text{ est indépendante de } N_u, \forall u \leq t.$$

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

$$- \text{La loi de } N_{t+s} - N_s \text{ ne dépend pas de } u.$$

1.4 Processus de Poisson

Beaucoup de phénomènes naturels peuvent présenter des changements de valeurs à n'importe quel moment plutôt qu'à des dates fixes. On aura besoin pour modéliser cela de processus en temps continu c.à.d. des processus $\{X_t : t \geq 0\}$ indicés par la demi droite positive $[0, +\infty)$ Les processus que l'on analyse ici sont à valeurs entières.

Le processus de Poisson sert à modéliser l'occurrence d'événements successifs. Chaque événement 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,
- les temps d'occurrence de sinistres à dédommager par une compagnie d'assurance...etc.

Définition 1.4.1. (Processus de poisson)

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

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

1.4.1 Loi de Poisson et loi exponentielle**Définition 1.4.1.1. (Loi exponentielle)**

Une variable aléatoire X de loi exponentielle de paramètre (ou de taux) $\lambda > 0$ est une variable continue à valeurs positives de densité

$$f(x) = \lambda e^{-\lambda x} \mathbf{1}_{x \geq 0}$$

Propriété 1.4.1.1. La loi exponentielle de paramètre λ est notée $\mathcal{E}(\lambda)$.

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

- Sa fonction de répartition est

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0. \end{cases}$$

- Sa fonction génératrice des moments est

$$\varphi(t) = \mathbb{E}[e^{tX}] = \begin{cases} \infty & \text{si } t \geq \lambda \\ \frac{\lambda}{\lambda - t} & \text{si } t < \lambda \end{cases}$$

- Sa moyenne et sa variance sont

$$\mathbb{E}[X] = \frac{1}{\lambda} \quad \text{Var}(X) = \frac{1}{\lambda^2}$$

Preuve

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

1. $F(t) = \mathbb{P}(X \leq t) = 0$ si $t < 0$ car X est une variable positive et si $t \geq 0$ on a

$$F(t) = \int_0^t f_X(x) dx = \int_0^t \lambda e^{-\lambda x} dx = [-e^{-\lambda x}]_0^t = 1 - e^{-\lambda t}$$

2. Sa fonction génératrice des moments vérifie

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

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

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

d'où

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

Définition 1.4.1.2. (Loi de poisson)

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

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

La loi de Poisson de paramètre μ est notée $\mathcal{P}(\mu)$.

Propriété 1.4.1.2. Soit X une variable aléatoire de loi de Poisson $\mathcal{P}(\mu)$

– Sa fonction génératrice des moments est

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

– Sa moyenne et sa variance sont

$$\mathbb{E}[X] = \mu \quad \text{Var}(X) = \mu$$

Preuve

$$\varphi(t) = \sum_{k=0}^{\infty} e^{tk} \frac{\mu^k}{k!} e^{-\mu} = e^{-\mu} \sum_{k=0}^{\infty} \frac{(\mu e^t)^k}{k!} = e^{-\mu} e^{\mu e^t}$$

On calcule les dérivées de la fonction génératrice

$$\varphi'(t) = \mu e^t e^{\mu(e^t-1)} \quad \varphi''(t) = (1 + \mu e^t) \mu e^{\mu(e^t-1)}$$

D'où

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

1.4.1.1 Relation entre la distribution Exponentielle et la distribution de Poisson

La densité de probabilité d'une distribution exponentielle $f(t) = \lambda e^{-\lambda t}$ Supposons τ est exponentielle avec une espérance $\frac{1}{\lambda}$, et n est de Poisson de moyenne μ on a :

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

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

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

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

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

...

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

1.4.1.2 Propriété sans mémoire de la distribution exponentielle

Une variable aléatoire \mathbf{X} est dite sans mémoire (ou sans usure) si :

$$\forall s, t \geq 0$$

$$P(X > t + s / X > t) = P(X > s)$$

Si \mathbf{X} est la durée de vie d'un matériel quelconque l'équation précédente s'interprète de la manière suivante, sachant le matériel en état de bon fonctionnement au temps t , la loi de probabilité de sa durée de vie future est la même que celle de sa durée de vie initiale. En d'autres termes, le matériel ne s'use pas.

1.5 Processus de naissance et de mort

Définition 1.5.1. (*Processus de naissance et de mort*)

On peut réaliser un processus de naissance et de mort de la façon suivante :

- Les arrivées et les départs d'entités obéissent à des lois exponentielles de taux respectifs $\lambda(n)$ et $\mu(n)$.

- A l'aide d'hypothèse de régularité : deux évènements ne peuvent pas se produire en même temps, donc la probabilité que deux évènements se produisent dans un intervalle de temps dt est négligeable.
- Il y a une transition vers un état voisin, soit par l'arrivée d'un client (naissance), soit par le départ d'un client (mort).

Si $\pi_n(t)$ est la probabilité pour qu'il y a n clients dans le système à l'instant t , l'équation de Kolomogorov [18] s'écrit, pour $n > 0$

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

C'est-à-dire, en faisant tendre dt vers 0, pour $n > 0$:

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

De la même façon, on obtient pour $n = 0$.

$$\frac{d}{dt} \pi_0(t) = -\lambda_0 \pi_0(t) + \mu_1 \pi_1(t).$$

Le processus de Poisson est un cas particulier du processus de naissance et de mort pour lequel $\mu_n = 0$ et $\lambda_n = \lambda$ mais dans ce cas, il n'y a pas de régime stationnaire [5] Les équations différentielles s'écrivent alors :

$$\frac{d}{dt} \pi_0(t) = -\lambda_0 \pi_0(t),$$

d'où $\pi_0(t) = e^{-\lambda t}$

$$\frac{d}{dt} \pi_n(t) = -\lambda (\pi_n(t) - \pi_{n-1}(t))$$

dont la solution est $\pi_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$

1.5.1 Processus de naissance

Définition 1.5.1.1. (Processus de naissance) Le processus de naissance est la généralisation direct d'un processus de poisson lorsque le paramètre d'intensité λ dépend de l'état courant du processus, il va nous permettre d'introduire le concept "d'explosion"

Si la taille d'une population a une transition $n \rightarrow n + 1$ donc il correspond à une naissance

1.5.2 Processus de mort

Si la taille d'une population a une transition $n \rightarrow n - 1$ une mort

Définition 1.5.2.1. (*Processus de naissance et de mort*) C'est un cas particulier de chaîne de Markov [17] où seules les transitions d'un état à un état voisin sont permises, on s'intéresse au cas continu avec des taux de transition C'est le point de départ de la théorie des files d'attente. On introduit les données suivantes :

- 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 .

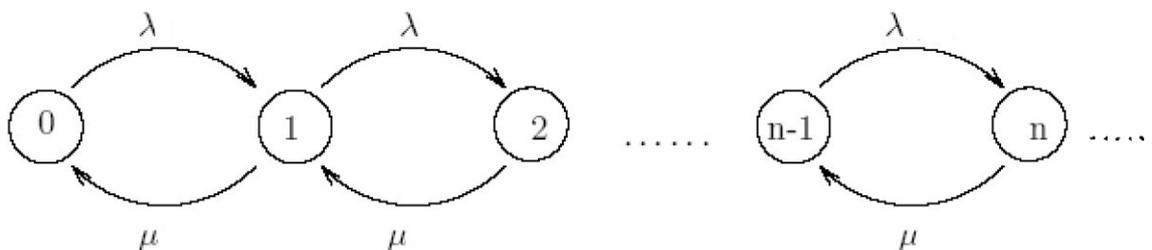


FIGURE 1.1 – Diagramme de transition d'un processus de naissance et de mort

Chapitre 2

Système de files d'attente

La théorie des files d'attente, ou queues, fournit des outils analytiques puissants pour la modélisation de systèmes admettant un phénomène d'attente. Cette théorie a pour objet l'étude des systèmes où des entités, appelées clients, cherchant à accéder à des ressources, afin d'obtenir un service. Les problèmes posés par l'économie où la gestion des entreprises, les applications dans les télécommunications, l'informatique et dans des domaines très divers ont été la cause principale de développement de cette théorie.

Grace à tous ces développements, la théorie des files d'attente est aujourd'hui largement utilisée et ses applications sont multiples.

Citons quelques exemples :

1. Les pannes de machines : l'élément est la machine qui tombe en panne, le poste est le mécanicien, le service est la réparation,
2. les arrivées de véhicules à un poste de chargement (arrivées de bateaux dans un port, d'avions au-dessus d'un aéroport, de bétonnières se charger en ciment, ...etc),
3. les arrivées de clients dans un magasin,
4. le règlement de la circulation au moyen des feux rouges et verts, ...etc.

2.1 Définitions

File d'attente : l'ensemble des clients qui attendent d'être servis, à l'exclusion de celui qui est en train de se faire servir.

Système d'attente : l'ensemble des clients qui font la queue, y compris celui qui se fait servir.

Le **phénomène d'attente** s'étend à tous les clients possibles (dans le cas de systèmes bouclés, où les mêmes clients reviennent plus tard à l'entrée par exemple les machines qui tombent en panne dans un atelier, le nombre des clients est, en général, fini). Ces

appellations se généralisent et prennent surtout leur intérêt dans les situations où il existe plusieurs stations et plusieurs files d'attente.

Classification des Systèmes d'attente [12]

Afin de spécifier un système de file d'attente, on donne les éléments suivants :

- La loi des intervalles séparant deux arrivées consécutives ou inter-arrivées ; c'est le processus d'entrée
- La loi du service : La loi du temps aléatoire qui met le client à recevoir le service.
- La discipline du service : Le nombre de stations de service disponible en même temps.
- La capacité totale du système N qui représente le nombre maximum de clients pouvant être présents dans le système

Remarque : cette dernière est supprimée si elle est infinie.

Notations et symboles [3]

- 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)
- $\bar{N} = \mathbb{E}(N)$: nombre moyen de clients dans le système,
- \bar{N}_S : nombre moyen de clients en train d'être servis,
- \bar{N}_Q : nombre moyen de clients dans la file d'attente, N_Q, N_S et N sont les v.a. correspondantes,
- \bar{T} : temps moyen qu'un client passe dans le système,
- \bar{T}_S : temps moyen de service,
- \bar{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 simple

File d'attente simple

Définition 2.2.1. (*File d'attente simple*) Une file simple (ou station) est un système constitué d'un ou plusieurs serveurs et d'un espace d'attente. Les clients arrivent de l'extérieur, patientent éventuellement dans la file d'attente, reçoivent un service, puis

quittent la station [9]. Afin de spécifier complètement une file simple, on doit caractériser le processus d'arrivé des clients, le temps de service ainsi que la structure et la discipline de service de la file d'attente.

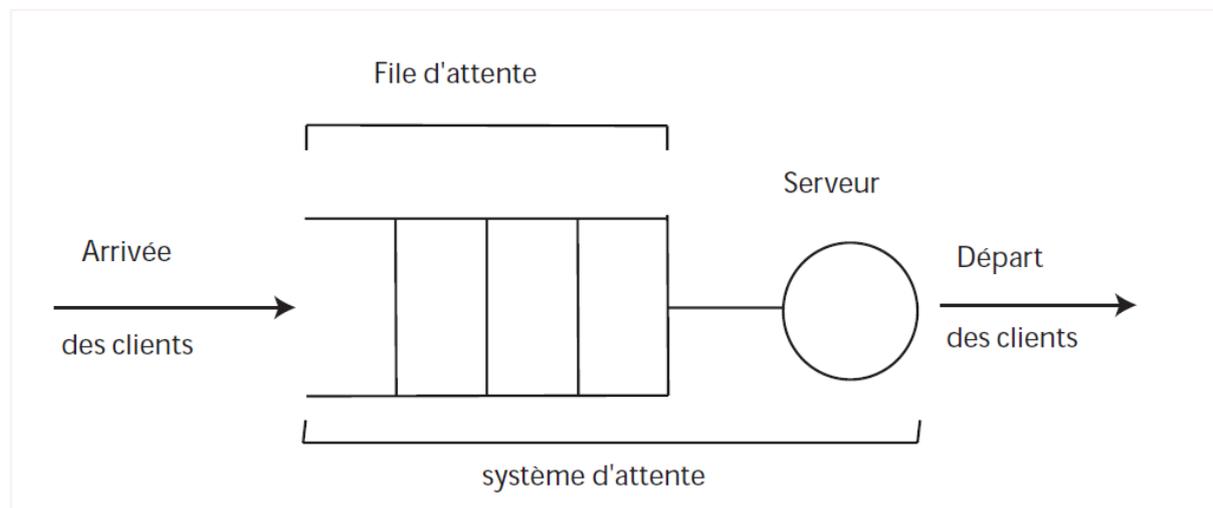


FIGURE 2.1 – le système de file d'attente

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 $n^{\text{ième}}$ client dans le système, on aura ainsi :

$$A_0 = 0 \quad \text{et} \quad A_n = \inf \{t; N_t = n\}$$

Si T_n désigne la variable aléatoire mesurant le temps séparant l'arrivée du $(n-1)^{\text{ième}}$ et du $n^{\text{ième}}$ client [10], on a alors :

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

2.2.2 Temps de service :

On considère une file à serveur unique.

On note D_n la variable aléatoire mesurant l'instant de départ du $n^{\text{ième}}$ client du système et Y_n la variable aléatoire mesurant le temps de service du $n^{\text{ième}}$ client (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 considèrera uniquement des stations dont les temps de service consécutifs sont décrits par des variables Y_n indépendantes et identiquement distribuées (i.i.d.).

Remarque 2.2.1. *La distribution du temps de service la plus simple à étudier est la distribution exponentielle. Cependant, la propriété sans mémoire" de la loi exponentielle fait que celle-ci n'est généralement pas très réaliste pour modéliser les phénomènes réels. On est donc souvent obligé de recourir à d'autres distributions de service.*

2.3 Structure et discipline de la file

2.3.1 Nombre de serveurs :

Le nombre de serveurs correspond au nombre maximal de clients pouvant être traités simultanément. Tous les serveurs sont supposés identiques, en particulier les temps de service sont indépendants d'un serveur à l'autre et distribués selon une même loi de probabilité.

2.3.2 Capacité de la file :

La capacité d'accueil d'un système de files d'attente correspond au nombre maximal de clients pouvant être présents dans le système à un instant quelconque. Il est égal à la somme du nombre de serveurs et du nombre de places d'attente disponibles. Si un client arrive dans un système ayant atteint sa capacité maximale d'accueil, il est refoulé et doit quitter le système sans avoir été servi. Le plus souvent, le nombre de clients susceptibles d'accéder au service est supposé illimité et leur fréquence d'arrivée constante. Certaines situations sont, cependant, caractérisées par un nombre fixe et limité de clients. Chaque client présent dans le système diminue alors le nombre d'arrivées potentielles, le taux d'arrivée dans la file n'est donc plus constant mais dépend du nombre de clients présents dans le système. Dans un tel cas, le processus d'arrivée décrit le temps nécessaire à un client entre le moment où il quitte le système et celui où il y revient.

2.3.3 Discipline de service :

La discipline d'attente, est la règle de priorité déterminant l'ordre dans lequel les clients vont accéder à la ressource modélisé par le serveur. Les disciplines d'attente classiques, ainsi que leurs acronymes, sont

- 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 la deuxième, il sera le premier à commencer son service. Rien n'empêche alors qu'un client qui commence son service après lui, dans un autre serveur, termine avant 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). À nouveau, les disciplines LIFO et LCFS ne sont équivalentes que pour une file mono serveur
- RANDOM (aléatoire) : le prochain client qui sera servi est choisi aléatoirement dans la file d'attente.
- Round-Robin (cyclique). Tous les clients de la file d'attente entrent en service à tour de rôle, effectuant un quantum Q de leur temps de service et sont remplacés dans la file, jusqu'à ce que leur service soit totalement accompli. Cette discipline de service a été introduite afin de modéliser les systèmes informatiques.
- PS (Processor Sharing), les clients sont servis de manière égale. La capacité du système est partagée entre les clients.

2.4 Notation de Kendall

Pour la classification des systèmes de files d'attente, on a recours à une notation symbolique introduite par Kendall [6], comprenant six symboles rangés dans l'ordre

$$A/B/s/N/D/O$$

avec

- 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 .
- B : décrit la distribution des temps de service d’un client. les codes sont les mêmes que A,
- s : nombre de serveurs,
- N : capacité de la file (c’est le nombre de places dans le système en d’autre tème c’est le nombre maximal de clients permis dans le système B compris ceux en service,
- D : population des usagers,
- O : discipline de service (c’est la façon dont les clients sont ordonnés pour être servi.

2.5 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 système. 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 La loi de Little s’exprime telle que dans la propriété suivante :

Théorème 2.5.1. (Formule de Little) : *Le nombre moyen de clients \bar{N} , le temps moyen passé dans le système \bar{T} et le débit moyen λ d’un système stable en régime permanent se relient de la façon suivante :*

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

où λ_e est le taux d’entrée dans le système ($\lambda_e = \lambda$ pour une file(M/M/1) 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).

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

La loi de Little peut aussi s'appliquer en considérant uniquement l'attente dans la queue (sans le service). Elle permet alors de relier le nombre moyen de clients en attente (N_Q) au temps moyen d'attente d'un client avant service (\bar{T}_Q)

par la relation : $\bar{N}_Q = \lambda_e \bar{T}_Q$

Enfin, on peut appliquer la loi de Little en ne considérant que le serveur. Dans ce cas, elle relie le nombre moyen de clients en service (\bar{N}_S), au temps moyen de séjour d'un client dans le serveur qui n'est rien d'autre que le temps moyen de service (\bar{T}_S) par la relation : $\bar{N}_S = \lambda_e \bar{T}_S$. On a obtenu trois relations en appliquant la loi de Little successivement au système entier à la file d'attente seule et enfin au serveur seul. Ces trois relations ne sont bien sûr pas indépendantes. On peut en effet déduire l'une d'entre elles à partir des deux autres en remarquant que $\bar{N} = \bar{N}_Q + \bar{N}_S$ et $\bar{T} = \bar{T}_Q + \bar{T}_S$

Remarque 2.5.1. *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).*

2.6 Mesure de performance d'une file d'attente

L'étude d'une file d'attente ou d'un réseau de files d'attente a pour but de calculer ou d'estimer les performances d'un système dans des conditions de fonctionnement données, et les mesures les plus fréquemment utilisées sont :

De manière générale, une file est stable si et seulement 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. Si chaque serveur peut traiter c clients par unité de temps et si le nombre de serveurs est m , une file est stable si et seulement si

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

où, ρ est appelé l'intensité du trafic.

2.7 Arrivée avant un départ et départ avant une arrivée

– Temps pour qu'une nouvelle arrivée se produise :

$$A \sim \mathcal{E}(\lambda)$$

– Temps pour qu'un nouveau départ se produise :

$$D \sim \mathcal{E}(\mu)$$

(A et D sont indépendants)

– Probabilité qu'une arrivée se produise avant un départ :

$$\mathbb{P}(A < D) = \frac{\lambda}{\lambda + \mu}$$

– Probabilité qu'un départ se produise avant une arrivée :

$$\mathbb{P}(D < A) = \frac{\mu}{\lambda + \mu}$$

2.8 Analyse en régime stationnaire

Il est difficile d'étudier la variable aléatoire $N(t)$ représentant le nombre de clients au temps t dans le système. On s'intéresse plutôt à $N = \lim_{t \rightarrow \infty} N(t)$, on parle alors d'analyse en régime stationnaire (ou analyse à l'équilibre). Pour qu'une file $M/M/1$ puisse atteindre l'équilibre par exemple, il faut que $\lambda < \mu$ (sinon la taille de la file augmentera à l'infini). À l'équilibre, on peut montrer que

$$\mathbb{P}(N = n) = \frac{\lambda}{\lambda + \mu} \mathbb{P}(N = n - 1) + \frac{\mu}{\lambda + \mu} \mathbb{P}(N = n + 1)$$

Il s'agit de la règle des probabilités totales. Le terme $\frac{\lambda}{\lambda + \mu}$ représente la probabilité qu'un nouveau client arrive avant que le client en service quitte le système, et $\frac{\mu}{\lambda + \mu} \mathbb{P}(N = n + 1)$ est la probabilité que le client en service quitte avant qu'un nouveau client n'arrive.

2.9 La file $M/M/1$

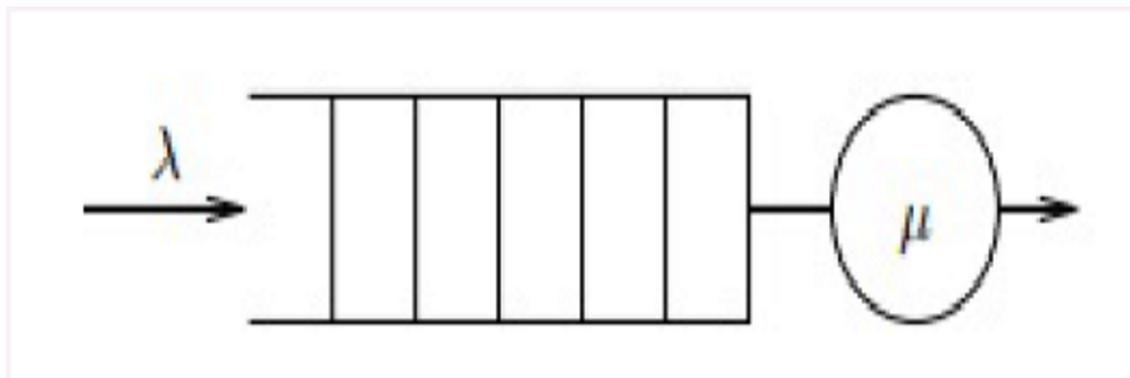
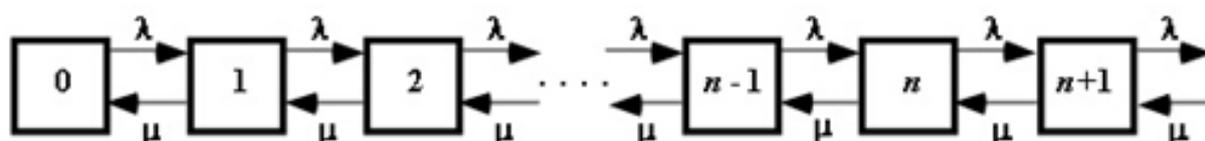
Le système de files d'attente $M/M/1$ est le système le plus élémentaire de la théorie des files d'attente. Le flot des arrivées est poissonnier de paramètre λ et la durée de service est exponentielle de paramètre μ , la discipline d'attente est FIFO la file d'attente est de capacité infinie.

La file peut être considérée comme un processus de naissance et de mort pour lequel figure (2.3)

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

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

Régime transitoire :

FIGURE 2.2 – La file $M/M/1$ FIGURE 2.3 – Diagramme de transition de la file d'attente $M/M/1$.

Soit $N(t)$ le nombre de clients présents dans le système à l'instant t ($t \geq 0$) grâce aux propriétés fondamentales du processus de Poisson et de la loi exponentielle, $N(t)$ est un processus markovien homogène.

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.

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

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

Régime stationnaire :

Sous la condition d'ergodicité du système $\rho = \frac{\lambda}{\mu} < 1$, pour laquelle le régime stationnaire existe, il est aisé d'obtenir les probabilités stationnaires

$$\begin{aligned} \pi_n &= \lim_{t \rightarrow \infty} p_n(t) \\ &= (1 - \rho)\rho^n, \quad \forall n \in \mathbb{N} \end{aligned}$$

$\pi = \{\pi_n\}_{n \geq 0}$ est appelé distribution stationnaire, elle suit une loi géométrique.

Caractéristiques du système :

- Le nombre moyen de clients dans le système est :

$$\begin{aligned}\bar{N} &= \mathbb{E}(N) \\ &= \sum_{n \geq 0} n\pi_n \\ &= (1 - \rho) \sum_{n \geq 0} n\rho^n\end{aligned}$$

D'où :

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

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

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

- Le nombre moyen de clients dans la file

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

Le temps moyen qu'un client passe dans le système T , le temps moyen de service \bar{T}_S et le temps moyen d'attente dans la file \bar{T}_Q sont obtenus à partir des formules de Little, ou des distributions du système

- Temps moyen qu'un client passe dans le système

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

- Temps moyen de service

$$\bar{T}_S = 1/\mu \tag{2.2}$$

- Temps moyen d'attente

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

2.10 La file $M/M/1/K$

On considère un système à serveur simple identique à la file $M/M/1$ excepté que la capacité de la file d'attente est finie. On a donc toujours les hypothèses suivantes : le processus d'arrivée des clients dans la file est un processus de Poisson de taux λ et le temps de service d'un client est une variable aléatoire exponentielle de taux μ . 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 : $E = \{0, 1, 2, \dots\}$ 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 "partir" 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'entrée, ce qui correspond bien à la stabilité inconditionnelle du système.

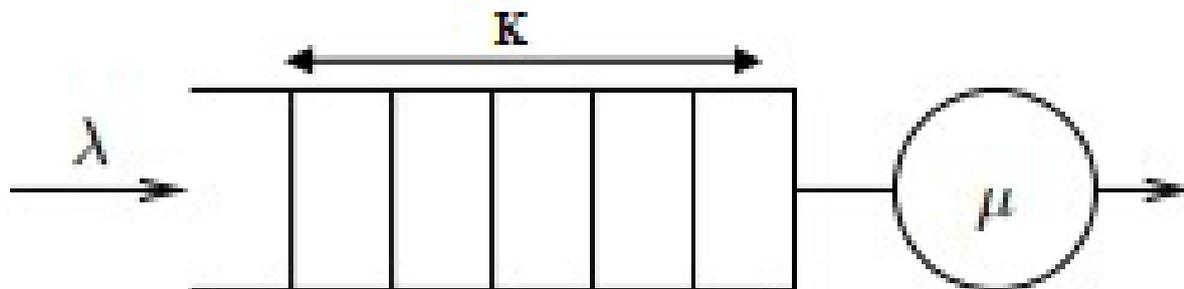


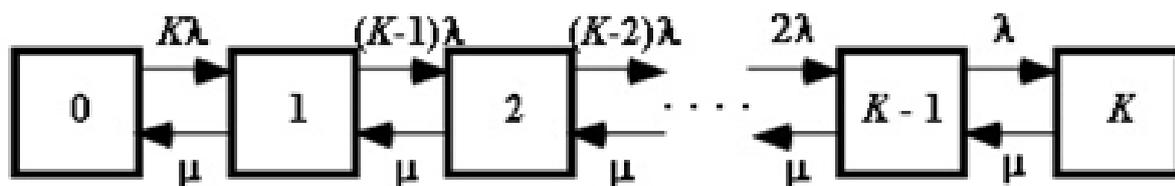
FIGURE 2.4 – La file $M/M/1/K$

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

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

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

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

FIGURE 2.5 – Evaluation de l'état dans la file d'attente $M/M/1/K$

et

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

Caractéristiques du système :

– Le nombre moyen de clients dans le système est :

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

À nouveau, lorsque K tend vers l'infini et $\rho < 1$, on retrouve les résultats de la file $M/M/1$:

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

– Le nombre moyen de clients dans la file est :

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

Le temps moyen qu'un client passe dans le système \bar{T} et le temps moyen d'attente dans la file \bar{T}_Q sont obtenus à partir la loi de Little :

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

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

– Temps moyen d'attente

$$\bar{T}_Q = \frac{\bar{N}_Q}{\lambda} \quad (2.4)$$

2.11 La 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 : processus d'arrivée des clients poissonien de taux λ et temps de service exponentiel de taux μ , (pour chacun des serveurs). 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 : $E = \{0, 1, 2, \dots\}$. 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

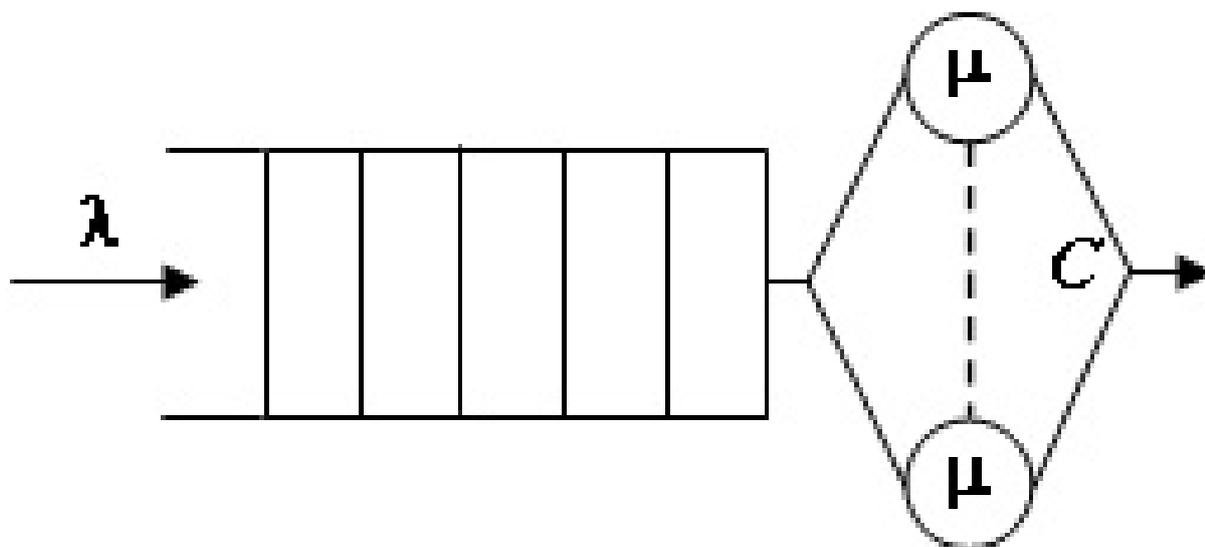
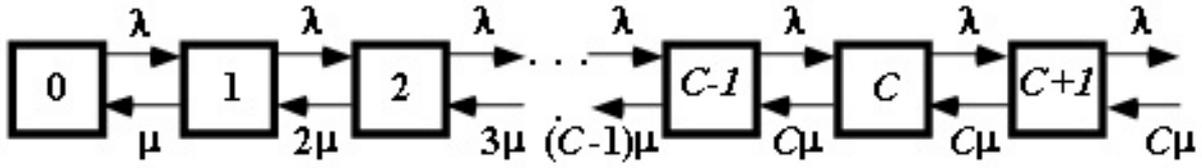


FIGURE 2.6 – La file $M/M/C$

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

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

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

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

Du diagramme, on déduit les résultats qui suivent. L'analyse du système en régime stationnaire, à l'aide de la procédure des équations de Chapman Kolmogorov aboutit aux équations suivantes :

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

avec

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

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

$$\begin{aligned}\bar{N}_Q &= \sum_{n=1}^{\infty} (n-1)\pi_n \\ &= \bar{N} - (1 - \pi_0) \\ \pi_n &= \frac{\rho^C}{C!} (A)^{n-C} \pi_0, \quad n \geq C\end{aligned} \tag{2.5}$$

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 : $\lambda < C\mu$

Caractéristiques du système : 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 est :

$$\bar{N} = \rho + \frac{\rho^{C+1}}{C \cdot C!(1-A)^2} \rho_0 \quad (2.6)$$

– Le nombre moyen de clients dans la file est :

$$\bar{N}_Q = \frac{\rho^{C+1}}{C \cdot C!(1-A)^2} \rho_0 \quad (2.7)$$

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

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

– Temps moyen d'attente :

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

2.12 La file $M/M/\infty$

On considère un système composé d'un nombre illimité de serveurs identiques et indépendants les uns des autres. Dès qu'un client arrive, il rentre donc instantanément en service. Donc cette file particulière, il n'y a donc pas d'attente. On suppose toujours que le processus d'arrivée des clients est poissonien de taux λ et que les temps de service sont exponentiels de taux μ (pour tous les serveurs). 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.

De façon intuitive, la capacité de traitement de la file est infinie puisque tout nouveau client se présentant à l'entrée de la file est instantanément traité. La condition de stabilité exprimant que "le nombre moyen de client arrivant à la file par unité de temps doit être inférieure à la capacité de traitement de la file" est donc toujours satisfaite.

Soit π_n la probabilité stationnaire d'être dans l'état n . Les équations d'équilibre nous donnent

$$\begin{aligned} \pi_{n-1}\lambda &= \pi_n n\mu \quad \text{pour } n = 1, 2, \dots \\ \text{soit } \pi_n &= \frac{\rho}{n}\pi_{n-1} \quad \text{pour } n = 1, 2, \dots, \end{aligned}$$

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

On peut alors exprimer toutes les probabilités en fonction de π_n .

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

La condition de normalisation nous donne alors immédiatement π_n

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

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

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

Caractéristiques du système :

– Nombre moyen de clients \bar{N} :

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

– Temps moyen de séjour \bar{T} :

Intuitivement, le temps moyen passé dans le système est réduit au temps moyen de service, soit $\frac{1}{\mu}$. On peut redémontrer ce résultat en utilisant la loi de Little :

$$\begin{aligned}\bar{T} &= \frac{\bar{N}}{\lambda_0} \\ &= \frac{1}{\mu}\end{aligned}$$

Chapitre 3

Système de files d'attente avec vacances de serveur

Dans un modèle de file d'attente classique, les serveurs sont toujours disponibles. Toutefois, dans de nombreux systèmes d'application pratiques, 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, par exemple en cours de maintenance ou simplement une pause. 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. Le fait de laisser les serveurs prendre des vacances rend les modèles de file d'attente plus réalistes et plus flexibles lors de l'étude des systèmes de files d'attente.

Les systèmes de file d'attente avec des vacances sur serveur ont attiré l'attention de nombreux chercheurs depuis que l'idée a été discutée pour la première fois dans l'article de Levy et Yechiali [16]. Plusieurs enquêtes sur ces modèles de vacances ont été réalisées par Doshi [7], [8] et les livres de Takagi [14], Tian et Zhang [15] sont consacrés à ce sujet.

3.1 Système de files d'attente avec vacance unique et vacances multiples

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. Le temps passé loin du centre de service principal s'appelle des 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 libérer 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.

3.1.1 Système de files d'attente avec vacance unique

Vacances unique : à la fin de chaque période de service, le serveur passe en vacances et revient immédiatement après la fin des vacances, même si le système est vide. Dans ce cas là il devient inactif jusqu'à l'arrivée d'un client. Le client est servi dès son arrivée. Par exemple, la maintenance des machines dans un processus de production est considérée comme une vacance unique.

3.1.2 Système de files d'attente avec vacances multiples

Vacances multiples : le serveur prend des vacances à chaque fois que le système devient vide. Si le serveur revient de vacances et trouve un système non vide, il démarre immédiatement le service et continue jusqu'à ce que le système devienne vide (service exhaustif). Si le serveur revient de vacances pour retrouver un système vide, il commence immédiatement une autre vacances et continue jusqu'à ce qu'il trouve un ou plusieurs clients à son retour.

Il existe différents types de systèmes de files d'attente avec vacances, selon la politique de service :

- La politique de service exhaustif, le serveur servira tous les clients en attente ainsi que ceux qui arrivent pendant qu'il sert encore. Il prend une autre vacance lorsque la file d'attente est vide.
- La politique de service bloqué, le serveur ne servira que les clients qu'il trouvera dans la file d'attente à son retour de vacances. À la fin de leur service, le serveur commencera une autre vacance et tous les clients qui arriveraient alors que le serveur servait déjà à la station seront servis au retour du serveur.
- La stratégie de service limité, le serveur ne servira qu'un nombre maximum prédéfini de clients, puis commencera par une autre vacance. Le programme de service de vacance unique dans lequel un seul client est servi est un type spécial de cette stratégie.

3.2 Système de files d'attente avec deux vacances différentes

3.2.1 Description mathématique du modèle

Nous considérons un système de files d'attente avec multiple vacances où les clients arrivent selon un processus de Poisson de paramètre λ . Le temps de service est supposé être distribué de manière exponentielle avec une moyenne de $1/\mu$, où $\mu > \lambda$. Nous supposons qu'il existe deux types de vacances : les vacances de type 1 prises après une période d'occupation de durée différente de zéro et les vacances de type 2 prises lorsqu'aucun client n'attend le serveur à son retour de vacances. Pour faciliter l'analyse, nous supposons que les durées des vacances de type 1 sont indépendantes de la période de pointe et sont distribuées de manière exponentielle avec une moyenne de $1/\gamma_1$ (Comme indiqué précédemment, il existe des cas où cette hypothèse d'indépendance n'est pas valide. Cependant, nous formulons cette hypothèse pour simplifier l'analyse.) De même, les durées de vacances de type 2 sont supposées être distribuées de manière exponentielle avec une moyenne de $1/\gamma_2$. Notons l'état du système par (r, k) , où r est le nombre de clients dans le système, $k = 0$ si le serveur est actif, servant des clients, $k = 1$ si le serveur est en vacances de type 1, et $k = 2$ si le serveur est en vacances de type 2. Ainsi, le système peut être modélisé par une chaîne de Markov à durée constante dont le diagramme de taux de transition d'état est montré à la figure (3.1).

3.2.2 Analyse stationnaire du modèle

D'après cette description, l'état du système à un instant arbitraire t peut être décrit par le processus stochastique à temps continu $\{(r(t), k(t)), t \geq 0\}$, où

- $r(t)$: est le nombre de clients dans le système à l'instant t ;
- $k(t)$: est l'état du serveur à l'instant t , avec

$$k(t) = \begin{cases} 0, & \text{si le serveur est occupé. ;} \\ 1, & \text{si le serveur est dans la période du vacance de type 1 ;} \\ 2, & \text{si le serveur est dans la période du vacance de type 2.} \end{cases}$$

Clairement, le processus $\{(r(t), k(t)), t \geq 0\}$ est un processus de Markov à temps continu et à valeurs dans l'espace d'état Ω tel que

$$\Omega = \{(n, 0) : n \geq 1\} \cup \{(n, j) : n \geq 0, j = 1, 2\}.$$

Maintenant, on définit en régime stationnaire :

– la probabilité qu’il y ait n clients dans le système à l’instant t et le serveur est occupé

$$P_{n,0} = \lim_{t \rightarrow +\infty} \mathbb{P} \left\{ r(t) = n, k(t) = 0 \right\}, n \geq 1,$$

– la probabilité qu’il y ait n clients dans le système à l’instant t et le serveur est en vacance de type 1

$$P_{n,1} = \lim_{t \rightarrow +\infty} \mathbb{P} \left\{ r(t) = n, k(t) = 1 \right\}, n \geq 0,$$

– la probabilité qu’il y ait n clients dans le système à l’instant t et le serveur est en vacance de type 2

$$P_{n,2} = \lim_{t \rightarrow +\infty} \mathbb{P} \left\{ r(t) = n, k(t) = 2 \right\}, n \geq 0.$$

Le diagramme de transition de ce processus est présenté dans la figure (3.1).

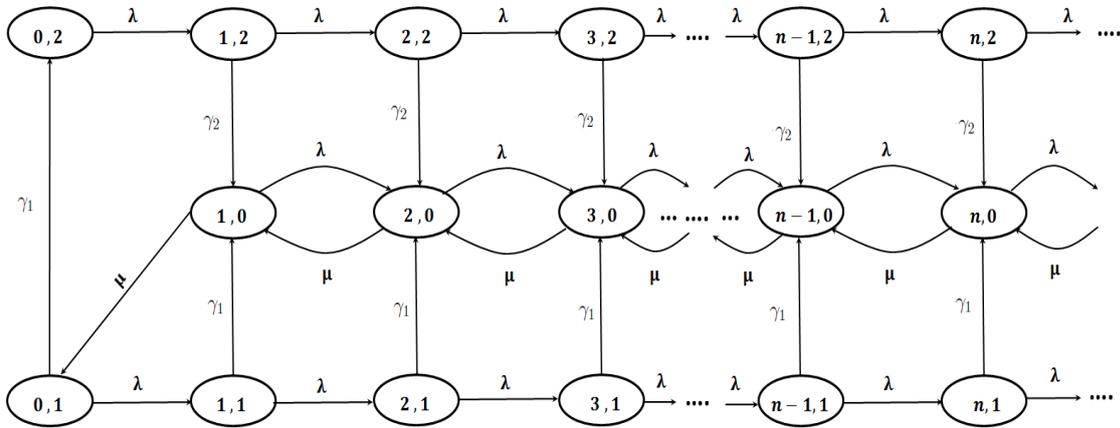


FIGURE 3.1 – le diagramme de transition de processus $\{(r(t), k(t))\}$.

Le diagramme voir figure (3.1) et l’analyse du système en régime stationnaire, à l’aide du procédé des équations de Chapman-Kolmogorov aboutissent aux équations d’équilibre suivantes :

$$(\lambda + \gamma_1) P_{0,1} = \mu P_{1,0} \tag{3.1}$$

$$\lambda P_{n,1} = (\lambda + \gamma_1) P_{n+1,1} \tag{3.2}$$

$$\lambda P_{0,2} = \gamma_1 P_{0,1} \tag{3.3}$$

$$\lambda P_{n,2} = (\lambda + \gamma_2) P_{n+1,2} \quad (3.4)$$

$$\lambda P_{n,0} + \lambda P_{n,1} + \lambda P_{n,2} = \mu P_{n+1,0} \quad n = 1, 2, \dots \quad (3.5)$$

avec la condition de normalisation

$$\sum_{n=1}^{\infty} P_{n,0} + \sum_{n=0}^{\infty} P_{n,1} + \sum_{n=0}^{\infty} P_{n,2} = 1 \quad (3.6)$$

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

Théorème 3.2.1. [13] *les probabilités de l'état du système en régime stationnaire $P_{n,k}$ sont données par :*

$$P_{n,k} = \begin{cases} \rho \left[\frac{\alpha_1 \beta_1 (\beta_1^{n-1} - \rho^{n-1})}{\beta_1 - \rho} + \frac{\alpha_2 \beta_2 (\beta_2^{n-1} - \rho^{n-1})}{\beta_2 - \rho} + \rho^{n-2} \right] P_{1,0} & k = 0 \\ \alpha_1 \beta_1^n P_{1,0} & k = 1 \\ \alpha_2 \beta_2^n P_{1,0} & k = 2 \end{cases} \quad (3.7)$$

avec

$$P_{1,0} = \left((1 - \rho) (1 - \beta_1) (1 - \beta_2) \right) \left((1 - \beta_1) (1 - \beta_2) + \alpha_1 (1 - \beta_2) \left\{ 1 - \rho (1 - \beta_1) \right\} + \alpha_2 (1 - \beta_1) \left\{ 1 - \rho (1 - \beta_2) \right\} \right)^{-1}. \quad (3.8)$$

Où

$$\rho = \lambda / \mu.$$

La quantité ρ représente l'intensité du trafic de système.

De plus,

$$\alpha_1 = \mu / (\lambda + \gamma_1), \quad \alpha_2 = \mu \gamma_1 / \lambda (\lambda + \gamma_1),$$

$$\beta_1 = \lambda / (\lambda + \gamma_1), \quad \beta_2 = \lambda / (\lambda + \gamma_2).$$

Preuve du théorème 3.2.1

L'équation (3.1), permet d'écrire,

$$P_{0,1} = \frac{\mu}{\lambda + \gamma_1} P_{1,0} = \alpha_1 P_{1,0} \quad (3.9)$$

avec $\alpha_1 = \mu / (\lambda + \gamma_1)$.

et l'équation (3.2), donne,

$$P_{n+1,1} = \frac{\lambda}{\lambda + \gamma_1} P_{n,1} \quad n = 0, 1, 2, \dots \quad (3.10)$$

En utilisant l'équation (3.10) et "la récursivité" sur n , nous obtenons

$$\begin{aligned} P_{n,1} &= \left(\frac{\lambda}{\lambda + \gamma_1} \right)^n P_{0,1}, \\ &= \alpha_1 \left(\frac{\lambda}{\lambda + \gamma_1} \right)^n P_{1,0}, \\ &= \alpha_1 \beta_1^n P_{1,0}, \quad n = 0, 1, 2, \dots \end{aligned}$$

avec $\beta_1 = \lambda / (\lambda + \gamma_1)$.

En suivant la même démarche, sur l'équation (3.3), nous obtenons,

$$P_{0,2} = \frac{\gamma_1}{\lambda} P_{0,1} = \left(\frac{\gamma_1}{\lambda} \right) \left(\frac{\mu}{\lambda + \gamma_1} \right) P_{1,0} = \alpha_2 P_{1,0} \quad (3.11)$$

avec $\alpha_2 = \mu \gamma_1 / \lambda (\lambda + \gamma_1)$.

De plus, l'équation (3.4), donne,

$$P_{n+1,2} = \frac{\lambda}{\lambda + \gamma_2} P_{n,2} \quad n = 0, 1, 2, \dots$$

D'où

$$\begin{aligned} P_{n,2} &= \left(\frac{\lambda}{\lambda + \gamma_2} \right)^n P_{0,2} \\ &= \alpha_2 \left(\frac{\lambda}{\lambda + \gamma_2} \right)^n P_{1,0} \\ &= \alpha_2 \beta_2^n P_{1,0} \quad n = 0, 1, 2, \dots \end{aligned}$$

avec $\beta_2 = \lambda / (\lambda + \gamma_2)$.

Maintenant, en utilisant l'équation (3.5), nous obtenons,

$$\begin{aligned} P_{n+1,0} &= \rho P_{n,0} + \rho P_{n,1} + \rho P_{n,2} \\ &= \rho P_{n,0} + \rho \alpha_1 \beta_1^n P_{1,0} + \rho \alpha_2 \beta_2^n P_{1,0} \end{aligned}$$

la méthode récursive sur n , donne,

$$\begin{aligned} P_{2,0} &= \rho \{ \alpha_1 \beta_1 + \alpha_2 \beta_2 + 1 \} P_{1,0}, \\ &= \rho \left\{ \frac{\alpha_1 \beta_1 (\beta_1 - \rho)}{\beta_1 - \rho} + \frac{\alpha_2 \beta_2 (\beta_2 - \rho)}{\beta_2 - \rho} + \rho^0 \right\} P_{1,0}, \\ P_{3,0} &= \rho \{ \alpha_1 \beta_1^2 + \alpha_2 \beta_2^2 + \rho \alpha_1 \beta_1 + \rho \alpha_2 \beta_2 + \rho \} P_{1,0}, \\ &= \rho \left\{ \frac{\alpha_1 \beta_1 (\beta_1^2 - \rho^2)}{\beta_1 - \rho} + \frac{\alpha_2 \beta_2 (\beta_2^2 - \rho^2)}{\beta_2 - \rho} + \rho \right\} P_{1,0}, \\ P_{4,0} &= \rho \{ \alpha_1 \beta_1^3 + \alpha_2 \beta_2^3 + \rho \alpha_1 \beta_1^2 + \rho \alpha_2 \beta_2^2 + \rho^2 \alpha_1 \beta_1 + \rho^2 \alpha_2 \beta_2 + \rho^2 \} P_{1,0}, \\ &= \rho \left\{ \frac{\alpha_1 \beta_1 (\beta_1^3 - \rho^3)}{\beta_1 - \rho} + \frac{\alpha_2 \beta_2 (\beta_2^3 - \rho^3)}{\beta_2 - \rho} + \rho^2 \right\} P_{1,0}. \end{aligned}$$

D'une façon générale, nous obtenons,

$$P_{n,0} = \rho \left\{ \frac{\alpha_1 \beta_1 (\beta_1^{n-1} - \rho^{n-1})}{\beta_1 - \rho} + \frac{\alpha_2 \beta_2 (\beta_2^{n-1} - \rho^{n-1})}{\beta_2 - \rho} + \rho^{n-2} \right\} P_{1,0} \quad n = 1, 2, \dots$$

En utilisant la condition de normalisation,

$$\sum_{n=1}^{\infty} P_{n,0} + \sum_{n=0}^{\infty} P_{n,1} + \sum_{n=0}^{\infty} P_{n,2} = 1$$

on obtient finalement $P_{1,0}$:

$$\begin{aligned} P_{1,0} &= \left((1 - \rho) (1 - \beta_1) (1 - \beta_2) \right) \left((1 - \beta_1) (1 - \beta_2) \right. \\ &\quad \left. + \alpha_1 (1 - \beta_2) \left\{ 1 - \rho (1 - \beta_1) \right\} + \alpha_2 (1 - \beta_1) \left\{ 1 - \rho (1 - \beta_2) \right\} \right)^{-1}. \end{aligned}$$

3.2.3 Mesures de performance

Par conséquent, la distribution stationnaire du processus permet de calculer d'autres paramètres de performances du système telles que :

La probabilité que le serveur soit dans la période de vacance de type 1 (P_{V_1}) :

$$\begin{aligned} P_{V_1} &= \sum_{n=0}^{\infty} P_{n,1} \\ &= \sum_{n=0}^{\infty} \alpha_1 \beta_1^n P_{1,0} \\ &= \alpha_1 P_{1,0} \sum_{n=0}^{\infty} \beta_1^n \end{aligned}$$

Par conséquent, $\beta_1 < 1$

$$P_{V_1} = \frac{\alpha_1}{1 - \beta_1} P_{1,0}.$$

La probabilité que le serveur soit dans la période de vacance de type 2 (P_{V_2}) :

$$\begin{aligned} P_{V_2} &= \sum_{n=0}^{\infty} P_{n,2} \\ &= \sum_{n=0}^{\infty} \alpha_2 \beta_2^n P_{1,0} \\ &= \alpha_2 P_{1,0} \sum_{n=0}^{\infty} \beta_2^n \end{aligned}$$

Par conséquent, $\beta_2 < 1$

$$P_{V_2} = \frac{\alpha_2}{1 - \beta_2} P_{1,0}.$$

La probabilité que le serveur soit dans la période d'occupation (P_B) :

$$\begin{aligned}
P_B &= \sum_{n=1}^{\infty} P_{n,0}, \\
&= 1 - \sum_{n=0}^{\infty} P_{n,1} - \sum_{n=0}^{\infty} P_{n,2}, \\
&= 1 - P_{V_1} - P_{V_2}.
\end{aligned}$$

Nombre moyen de clients dans le système (L_s) :

$$\begin{aligned}
L_s &= \sum_{n=1}^{\infty} nP_{n,0} + \sum_{n=0}^{\infty} nP_{n,1} + \sum_{n=0}^{\infty} nP_{n,2}, \\
&= \left\{ \frac{\alpha_1\beta_1\rho(2-\rho-\beta_1)}{(1-\rho)^2(1-\beta_1)^2} + \frac{\alpha_2\beta_2\rho(2-\rho-\beta_2)}{(1-\rho)^2(1-\beta_2)^2} \right. \\
&\quad \left. + \frac{1}{(1-\rho)^2} + \frac{\alpha_1\beta_1}{(1-\beta_1)^2} + \frac{\alpha_2\beta_2}{(1-\beta_2)^2} \right\} P_{1,0}.
\end{aligned}$$

Nombre moyen de clients dans la file (L_q) :

$$\begin{aligned}
L_q &= \sum_{n=2}^{\infty} (n-1)P_{n,0} + \sum_{n=0}^{\infty} nP_{n,1} + \sum_{n=0}^{\infty} nP_{n,2}, \\
&= \sum_{n=1}^{\infty} nP_{n,0} - \sum_{n=1}^{\infty} P_{n,0} + \sum_{n=0}^{\infty} nP_{n,1} + \sum_{n=0}^{\infty} nP_{n,2}, \\
&= \sum_{n=1}^{\infty} nP_{n,0} + \sum_{n=0}^{\infty} nP_{n,1} + \sum_{n=0}^{\infty} nP_{n,2} - \sum_{n=1}^{\infty} P_{n,0}, \\
&= L_s - P_B.
\end{aligned}$$

D'après les formules de Little, on obtient

Temps moyen de séjour dans le système (W_s) :

$$W_s = \frac{L_s}{\lambda}.$$

Temps moyen de séjour dans la file (W_q) :

$$W_q = \frac{L_q}{\lambda}.$$

3.2.4 Application numérique

Dans le but d'illustrer et de confirmer les résultats obtenu dans (3.2.1) précédente, nous avons élaboré un programme, sous logiciel libre *R*. certains résultats numériques sont présentés sous forme de tableaux et d'autres sous forme de graphes.

3.2.4.1 L'influence de taux d'arrivées des clients λ sur les différents paramètres de performances

Afin d'analyser l'influence du taux d'arrivées des clients λ sur les mesures de performance du modèle comme P_B , P_{V_1} , P_{V_2} , L_s , L_q , W_s , W_q , nous avons varié λ entre 0.6 et 1.5 par pas de 0.1, et nous avons fixé les autres paramètres comme : $\mu = 2$, $\gamma_1 = 0.5$, $\gamma_2 = 0.8$.

Nous obtenons les résultats suivants (3.1) :

TABLE 3.1 – Variations des paramètres de performance vs. λ .

λ	P_B	P_{V_1}	P_{V_2}	L_s	L_q	W_s	W_q
0.6	0.300000	0.420957	0.279043	1.449186	1.149186	2.415311	1.915311
0.7	0.350000	0.417192	0.232808	1.750424	1.400424	2.500606	2.000606
0.8	0.400000	0.405195	0.194805	2.071861	1.671861	2.589827	2.089827
0.9	0.450000	0.386881	0.163119	2.417990	1.967990	2.686655	2.186655
1.0	0.500000	0.363636	0.136364	2.795455	2.295455	2.795455	2.295455
1.1	0.550000	0.336484	0.113516	3.214110	2.664110	2.921918	2.421918
1.2	0.600000	0.306191	0.093809	3.688931	3.088931	3.074109	2.574109
1.3	0.650000	0.273342	0.076658	4.243595	3.593595	3.264304	2.764304
1.4	0.700000	0.238387	0.061613	4.917687	4.217687	3.512634	3.012634
1.5	0.750000	0.201681	0.048319	5.782563	5.032563	3.855042	3.355042

Pour ce premier tableau (3.1) : Par la croissance de λ , la probabilité d'occupation du serveur augmente ainsi que le nombre moyen de clients dans le système et dans la file et le temps moyen de séjour dans le système et dans la file ce qui logique. Par contre les probabilités que le serveur prend des vacances soit de type 1 soit de type 2 diminuent cela est interprété par le fait que le système ne se vide pas.

3.2.4.2 L'influence de taux service μ sur les différents paramètres de performances

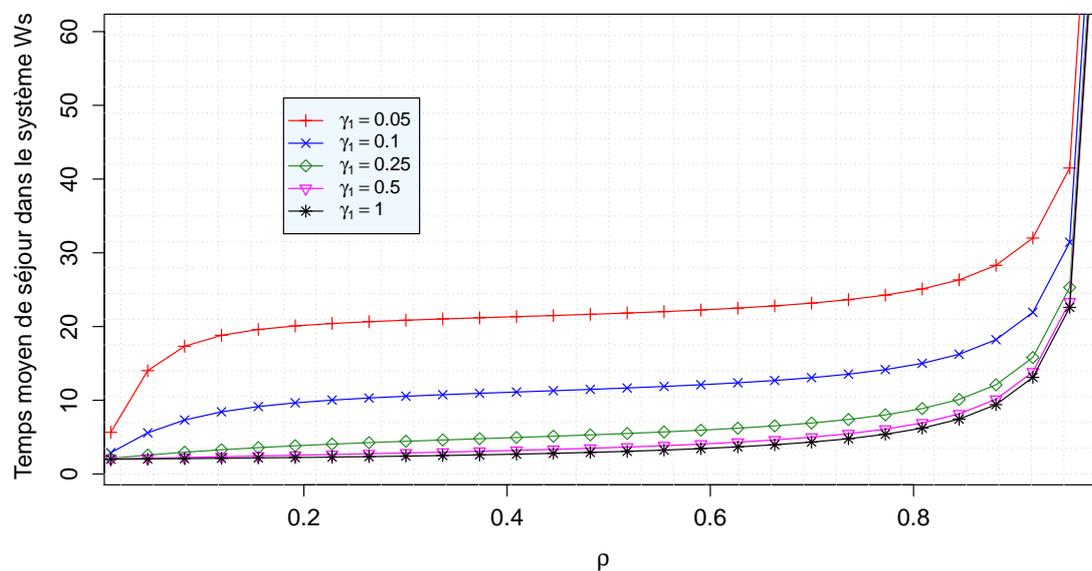
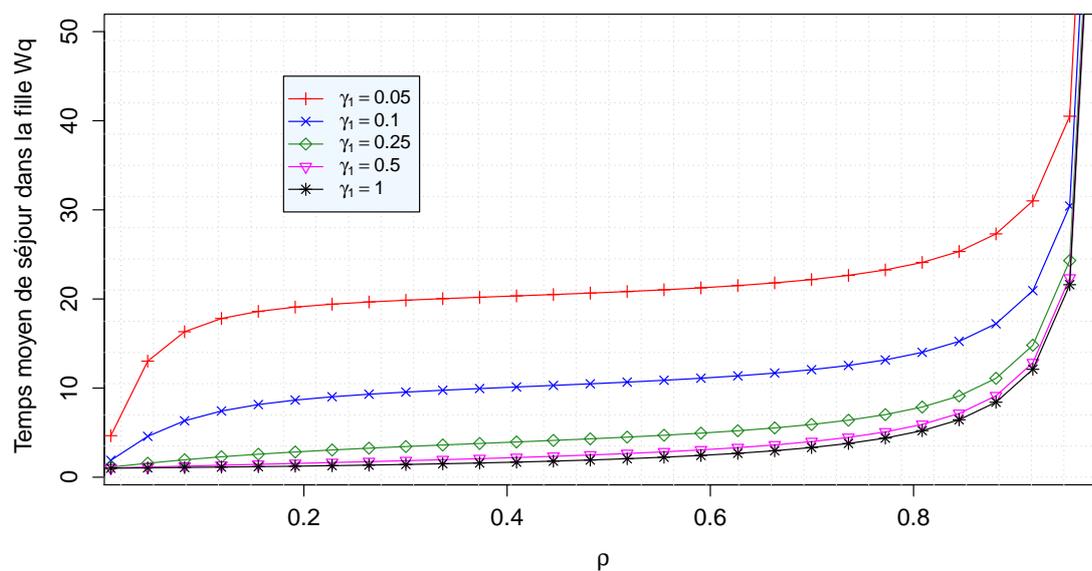
Afin d'analyser l'influence du taux de service des clients μ sur les mesures de performance du modèle comme P_B , P_{V_1} , P_{V_2} , L_s , L_q , W_s , W_q , nous avons varié les valeurs de μ entre 1.6 et 2.5 par pas de 0.1, et nous avons fixé les autres paramètres comme : $\lambda = 1.2$, $\gamma_1 = 0.5$, $\gamma_2 = 0.8$.

Nous obtenons les résultats suivants (3.2) :

TABLE 3.2 – Variations des paramètres de performance vs. μ .

μ	P_B	P_{V_1}	P_{V_2}	L_s	L_q	W_s	W_q
1.6	0.750000	0.191370	0.058630	5.188931	4.438931	4.324109	3.699109
1.7	0.705882	0.225141	0.068977	4.588931	3.883048	3.824109	3.235874
1.8	0.666667	0.255159	0.078174	4.188931	3.522264	3.490775	2.935220
1.9	0.631579	0.282018	0.086403	3.903216	3.271637	3.252680	2.726364
2.0	0.600000	0.306191	0.093809	3.688931	3.088931	3.074109	2.574109
2.1	0.571429	0.328062	0.100509	3.522264	2.950835	2.935220	2.459029
2.2	0.545455	0.347945	0.106601	3.388931	2.843476	2.824109	2.369563
2.3	0.521739	0.366098	0.112162	3.279840	2.758101	2.733200	2.298417
2.4	0.500000	0.382739	0.117261	3.188931	2.688931	2.657442	2.240775
2.5	0.480000	0.398049	0.121951	3.112008	2.632008	2.593340	2.193340

Pour ce tableau (3.2) : Par la croissance de μ , la probabilité d'occupation du serveur diminue ainsi que le nombre moyen de clients dans le système et dans la file et le temps moyen de séjour dans le système et dans la file ce qui logique. Par contre les probabilités que le serveur prend des vacances soit de type 1 soit de type 2 augmente celà est interprété par le fait que le système se vide.

3.2.4.3 L'influence des taux de vacances γ_1 sur (W_s) et (W_q) FIGURE 3.2 – Temps moyen de séjour dans le système (W_s) vs. ρ et γ_1 pour $\gamma_2 = 1$.FIGURE 3.3 – Temps moyen de séjour dans la file (W_q) vs. ρ et γ_1 pour $\gamma_2 = 1$.

Les résultats des figures (3.2), (3.3), montrent que plus ρ est grand le temps de séjour dans le système et dans la file sont grands. De plus le temps de séjour dans le système et

dans la file sont plus important dans le cas par exemple $\gamma_1 = 0.05$ (le taux de vacances de type 1) que celui de $\gamma_1 = 1$. Celà est interprété par le fait que $1/\gamma_1$ est le temps moyen de vacances de type 1 et par conséquent plus ce temps augmente plus le temps de séjours (attente) dans le système et dans la file augmentent.

3.2.4.4 L'influence des taux de vacances γ_2 sur (W_s) et (W_q)

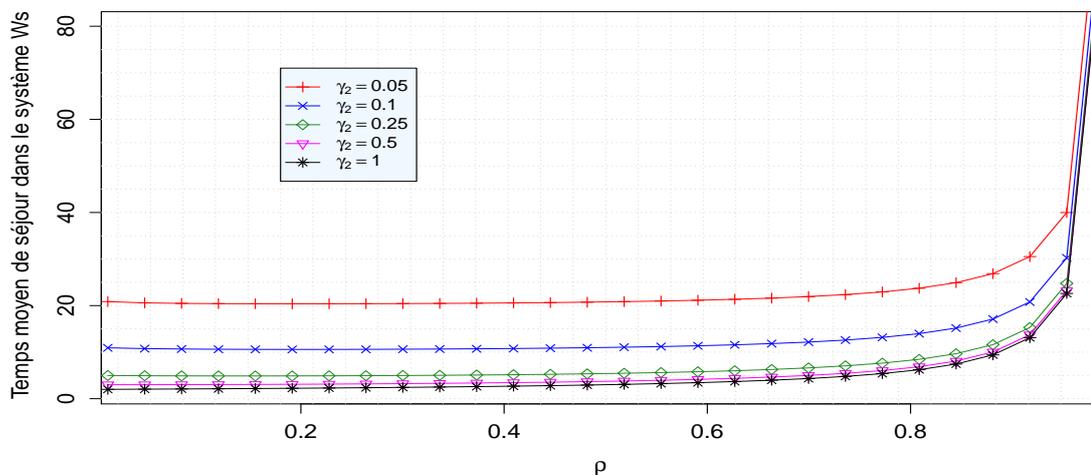


FIGURE 3.4 – Temps moyen de séjour dans le système (W_s) vs. ρ et γ_2 pour $\gamma_1 = 1$.

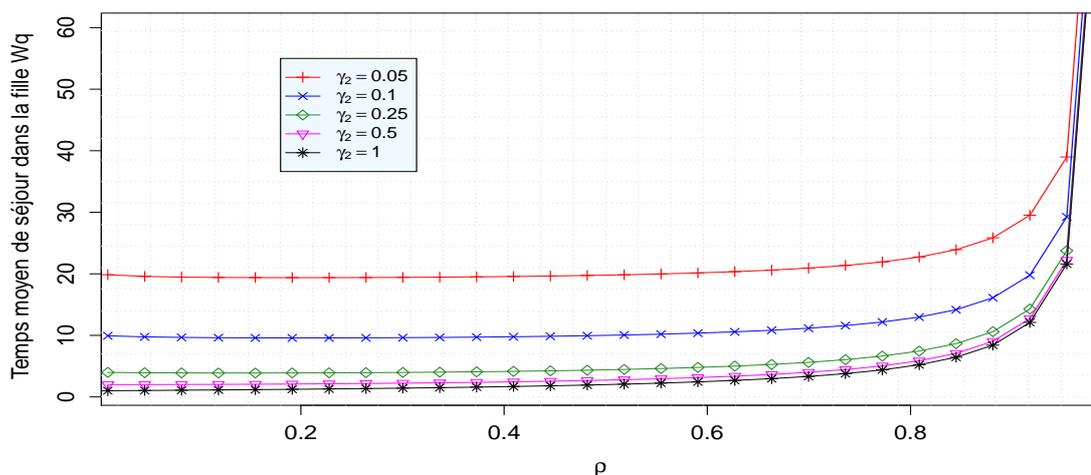


FIGURE 3.5 – Temps moyen de séjour dans la file (W_q) vs. ρ et γ_2 pour $\gamma_1 = 1$.

Pour les figures (3.4),(3.5), pareille que précédament, les résultats montrent que plus ρ est grand le temps de séjour dans le système et dans la file sont grands. De plus le temps

de séjour dans le système et dans la file sont plus important dans le cas $\gamma_2 = 0.05$ (le taux de vacances de type 2) que celui de $\gamma_2 = 1$. Celà est interprété par le fait que $1/\gamma_2$ est le temps moyen de vacances de type 2 et par conséquent plus ce temps augmente plus le temps de séjours (attente) dans le système et dans la file augmentent.

D'une manière générale, les résultats théoriques obtenus dans la (3.2.1) précédente, sont confirmés par les résultats numériques (une bonne concordance entre les résultats analytiques et ceux issus de l'application numériques).

Conclusion

Dans ce mémoire, nous avons traité un article de Oliver C. Ibe and Olubukola A. Isijola Publié le 23 July 2014. Ils ont considéré un système de files d'attente multiples dans lesquels deux types de vacances sont rencontrés. Le premier type correspond aux vacances simples prises à la fin d'une période d'occupation de durée différente de zéro, et le second correspond aux vacances multiples. Le cas $M/M/1$ est considéré. Ce système de files d'attente avec vacances a de nombreuses applications de la vie réelle dans lesquelles certaines vacances peuvent être utilisées pour des activités de poste traitement, tandis que d'autres sont de véritables "pauses" prises par le serveur.

Enfin une étude numérique sous logiciel R est donnée pour voir l'influence des différents paramètres du système les uns sur les autres, et sur les paramètres de performances du système .

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] Allen, A. O. 1990. Probability, Statistics, and Queueing Theory with Computer Science Applications. Second edition, Academic Press, New York (First edition :1978).
- [3] Anisimov, V. V., Zakusilo, O. K., and Donchenko, V. S. 1987. Elements of Queueing Theory and Asymptotic System Analysis. Vishcha Shkola, Kiev (in Russian).
- [4] Chaudhry, M. L., and Templeton, J. G. C. A First Course in Bulk Queues. John Wiley and Sons, New York. (1983).
- [5] Claudie Chabriac. Processus stochastiques et modélisation. (2012-2013).
- [6] Claudie Hasseforder CHABRIAC , Eléments de Théorie des files d'attente, page05, Janvier 2008.
- [7] B. T. Doshi, "Queueing systems with vacations-a survey," Queueing Systems : Theory and Applications, vol. 1, no. 1, pp. 29-66, 1986.
- [8] B. Doshi, "Single server queues with vacations," in Stochastic Analysis of Computer and Communication Systems, H. Takag, Ed., pp. 217-265, Elsevier, 1990.
- [9] K.B, GK, Techniques de modélisation : Méthodes analytiques.
- [10] 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.
- [11] LeGall, P. (1962). Les systemes avec ou sans attente et les processus stochastiques, Tome I. Dunod, Paris.
- [12] Newell, G. F. 1982. Applications of Queueing Theory. Second edition, Chapman and Hall, London (First edition : 1971).

-
- [13] Oliver C. Ibe and Olubukola A. Isijola. Published 23 July 2014. *M/M/1 Multiple Vacation Queueing Systems with Differentiated Vacations*, Hindawi Publishing Corporation Modelling and Simulation in Engineering Volume 2014.
- [14] 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.
- [15] N. Tian and Z. G. Zhang, *Vacation Queueing Models : Theory and Applications*, Springer, New York, NY, USA, 2006
- [16] 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.
- [17] Yves Caumel. *Probabilités et processus stochastiques*. Springer. Verlag France, (2011).
- [18] Zakhar Kabluchko, *Stochastic Processes (Stochastik II)*, University of Ulm Institute of Stochastics, (2013-2014).