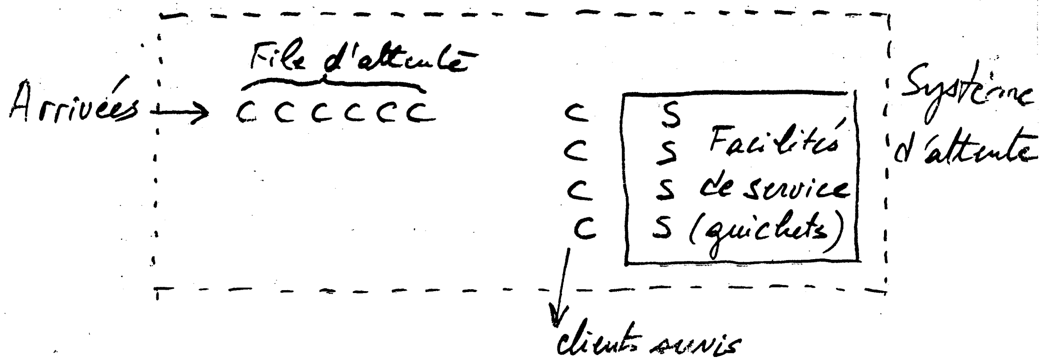
Chapitre 5 : Introduction à la théorie des files d'attente

(I) <u>Systèmes d'attente</u> : on est en présence d'une file d'attente quand des "clients" arrivent à des instants aléatoires dans un système et doivent attendre un certain temps avant de recevoir un service et pouvoir ensuite ressortir du système.

Les "<u>clients</u>" ne sont pas forcément des humains mais peuvent être des <u>entités quelconques</u> d'une même nature comme des dossiers, des appels téléphoniques, des informations numérisées etc...

On peut étudier un système d'attente quand on connait :

- les lois qui régissent les arrivées des clients
- le mécanisme de service (nombre de serveurs et lois de service)



<u>Paramètres descriptifs d'un système d'attente</u>

En vue d'une introduction élémentaire nous allons présenter seulement les quelques paramètres auxquels nous allons nous intéresser dans ce chapitre. La théorie générale nécessiterait la présentation préalable de la théorie des processus aléatoires en temps continu ce que nous voulons éviter, nous préférons introduire les processus de manière "ad-hoc" au fur et à mesure des besoins.

<u>Définitions et notations</u> :

1) On appelle <u>ligne d'attente</u> le nombre total de clients dans le système d'attente et <u>file d'attente</u> le nombre de clients qui attendent pour être servis.

2) E_n = l'événement "il y a n clients dans la ligne d'attente" (i.e. dans le système)

$P_n(t)$ = la probabilité qu'il y ait exactement n clients dans le système à l'instant t.

s = le nombre de serveurs (ou guichets) dans le système.

λ_n = le taux d'arrivées moyen (i.e. l'espérance du nombre d'arrivées par unité de temps) des nouveaux clients lorsque le système est dans l'état E_n

μ_n = le taux moyen de service (i.e. l'espéra-

-ce du nombre de clients servis par unité de temps) lorsque le système est dans l'état E_n.

On dit aussi que λ_n (resp. μ_n) est l'<u>intensité</u> d'arrivées (resp. de sorties) lorsque le système est dans l'état E_n.

<u>Cas particuliers:</u>

- si λ_n est constant ($\forall n$), on note $\lambda_n = \lambda$

De même, on note $\mu_n = \mu$ lorsque μ_n ne dépend pas de n. Dans ce cas, on a

$\frac{1}{\lambda}$ = le temps moyen entre 2 arrivées de clients

$\frac{1}{\mu}$ = le temps moyen de service d'un client

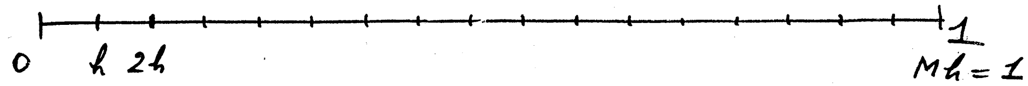
$\rho = \frac{\lambda}{s\mu}$ = le facteur d'utilisation des facilités de service (la moyenne de la proportion du temps pendant lequel les serveurs sont occupés)

II) <u>Le processus de Poisson</u>:

Le processus de Poisson est le plus simple de tous les processus de Markov en temps continu et à espace d'états discret. Il intervient dans un grand nombre nombre de situations concrètes. En particulier dans un système d'attente lorsque les clients arrivent <u>de manière purement aléatoire</u> à un taux moyen constant λ et qu'ils ne sont pas servis (i.e $s=0$).

Précisément expliquons ce que nous entendons par arrivées purement aléatoires:

Considérons un intervalle de temps unitaire (i.e. de longueur 1 par exemple 1 heure ou 1 minute) divisons cet intervalle en M parties égales de longueur $h = \frac{1}{M}$



<u>Hypothèses:</u>

1) Les arrivées qui se produisent dans les sous intervalles $[0,h[$, $[h,2h[$, ..., $[(M-1)h, 1[$ sont indépendantes (i.e. des v.a. indépendantes)
2) La probabilité d'une arrivée est la même quel que soit le sous-intervalle $[0,h[$, $[h,2h[$, ... (stationnarité des arrivées)

<u>Conséquences:</u>

- la probabilité qu'il y ait <u>au moins une arrivée</u> dans un sous-intervalle $[(k-1)h, kh[$ est égale à $1-P_0(h)$ $(\forall k)$
- l'espérance du nombre des intervalles où il y a au moins une arrivée est donc

$$M(1-P_0(h)) = \frac{1-P_0(h)}{h}$$

<u>Précision sur l'hypothèse que le taux moyen d'arrivées est constant et égal à $\lambda > 0$:</u>

$$\lim_{h \to 0} \frac{1 - P_0(h)}{h} = \lambda$$

<u>Hypothèse de rareté du nombre d'arrivées dans un intervalle de temps très petit :</u>

la probabilité qu'il y ait plus d'une arrivée dans un intervalle de temps de longueur h est négligeable devant h i.e. $o(h)$, autrement dit :

$$1 - P_0(h) - P_1(h) = o(h)$$

c'est à dire

$$\frac{1}{h}\left(1 - P_0(h) - P_1(h)\right) \longrightarrow 0 \quad \text{si } h \to 0$$

ou encore :

l'espérance du nombre d'intervalles de longueur h où il y a plus d'une arrivée, tend vers zéro si $h \to 0$.

<u>Etude de la fonction $P_n(t)$</u> : On peut <u>synthétiser</u> la discussion précédente par les 3 conditions suivantes sur le flux des arrivées :

H_1 : la probabilité $P_n(t)$ qu'il se produise n arrivées dans un intervalle de temps de longueur t dépend seulement de n et t et non de la position de l'intervalle sur l'axe des temps.

H2 : les nombres d'arrivées dans des intervalles de temps disjoints sont des v.a. indépendantes

H3 : la probabilité d'avoir plus d'une arrivée dans un (petit) intervalle de temps de longueur h est négligeable devant h (i.e. $o(h)$ quand $h \to 0$)

<u>Théorème :</u> Sous les hypothèses H1, H2 et H3, il existe une constante $\lambda > 0$ telle que $\forall t > 0$:

$$P_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \quad (n \in \mathbb{N})$$

Autrement dit le nombre d'arrivées dans un intervalle de temps de longueur t est une v.a. de Poisson de paramètre λt.

De plus λ = le taux moyen d'arrivées par unité de temps.

<u>démonstration :</u> l'hypothèse H 1 implique que $\forall t > 0$ et $h > 0$, on a :

$$P_0(t+h) = P_0(t)P_0(h).$$

De plus comme $t \mapsto P_0(t)$ est décroissante, c'est un exercice classique de montrer qu'il existe $\lambda > 0$ telle que

$$P_0(t) = e^{-\lambda t}.$$

Pour $h \to 0$, on a donc

$$P_0(h) = 1 - \lambda h + o(h).$$

Mais d'après H3 on a

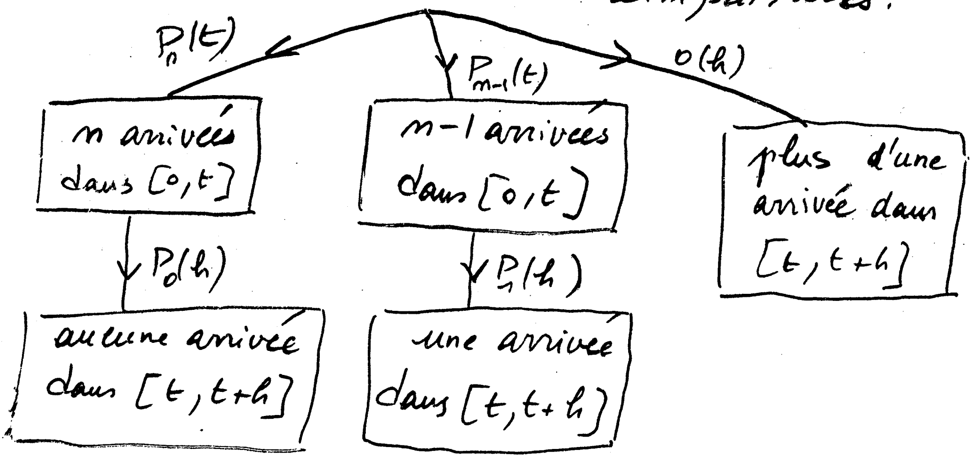
$$P_0(h) + P_1(h) + o(h) = 1,$$

d'où

$$P_1(h) = \lambda h + o(h).$$

Calculons maintenant $P_n(t+h)$:

L'événement "n arrivées se produisent dans l'intervalle $[0, t+h]$" peut se décomposer en réunion de 3 événements incompatibles:



D'où:

$$P_n(t+h) = P_0(h)P_n(t) + P_1(h)P_{n-1}(t) + o(h)$$

En remplaçant $P_0(h)$ et $P_1(h)$ par leurs valeurs (p.7), on a:

$$P_n(t+h) = (1-\lambda h)P_n(t) + \lambda h P_{n-1}(t) + o(h)$$

D'où

$$\frac{P_n(t+h) - P_n(t)}{h} = \lambda\left[P_{n-1}(t) - P_n(t)\right] + \frac{o(h)}{h}.$$

Si $h \to 0$, il vient:

$$P_n'(t) = \lambda\left[P_{n-1}(t) - P_n(t)\right] \quad (n \in \mathbb{N}^*).$$

En posant $\varphi_n(t) = e^{\lambda t}P_n(t)$, on obtient

$$\varphi_n'(t) = \lambda\,\varphi_{n-1}(t)$$

Avec les conditions initiales $\varphi_0(t) = 1$ et $\varphi_n(0) = 1$, on déduit:

$$\varphi_1(t) = \lambda t, \quad \varphi_2(t) = \frac{(\lambda t)^2}{2!}, \ldots$$

et par récurrence $\varphi_n(t) = \dfrac{(\lambda t)^n}{n!}$, d'où

$$P_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \quad \underline{\text{cqfd}}.$$

Pour tout $t > 0$, considérons la v.a.

X_t = nombre d'arrivées entre les instants 0 et t. Alors X_t est une v.a. de Poisson de paramètre λt. La famille des v.a. $(X_t)_{t\geq 0}$ est appelée <u>processus de Poisson</u> d'intensité $\lambda > 0$. C'est un processus de Markov à temps continu et à espace d'états discret ($= \mathbb{N}$)

<u>Instants d'arrivée dans un processus de Poisson</u>:

On a vu que

$$P_0(t) = \mathbb{P}(\text{pas d'arrivée dans } [0,t]) = e^{-\lambda t}.$$

Par conséquent

$$\mathbb{P}(1^{\text{ère}} \text{ arrivée ait lieu dans } [0,t]) = 1 - e^{-\lambda t}$$

Ainsi la variable aléatoire

$$T_1 = \text{instant de la } 1^{\text{ère}} \text{ arrivée}$$

a pour fonction de répartition

$$F(t) = \mathbb{P}(T_1 \leq t) = 1 - e^{-\lambda t} \quad (t > 0)$$

T_1 est donc une <u>v.a. exponentielle de paramètre $\lambda > 0$</u>. Plus généralement, on montre (exercice) que si T_n est l'instant de la $n$ième arrivée, les <u>temps d'attente</u> entre deux arrivées successives:

$$T_1,\ T_2 - T_1,\ T_3 - T_2, \dots, T_n - T_{n-1}, \dots$$

sont des v.a. indépendantes de même loi exponentielle de paramètre $\lambda > 0$

<u>Exemple</u> : le processus de <u>la désintégration naturelle</u> d'un élément radioactif (par exemple de l'uranium naturel) peut être modélisé par un processus de Poisson d'intensité $\lambda =$ le taux moyen de désintégration par unité de temps: si on identifie la désintégration d'un atome à l'arrivée d'un client, les hypothèses H1, H2 et H3 du Théorème sont généralement satisfaites. La probabilité qu'il y ait n atomes qui se désintègrent entre les instants 0 et t est égale à $e^{-\lambda t}\frac{(\lambda t)^n}{n!}$ et le nombre moyen d'atomes qui disparaissent entre les instants 0 et t est λt donc la masse moyenne qui disparait est proportionnelle à λt (comme on l'apprend dans les cours de physique).

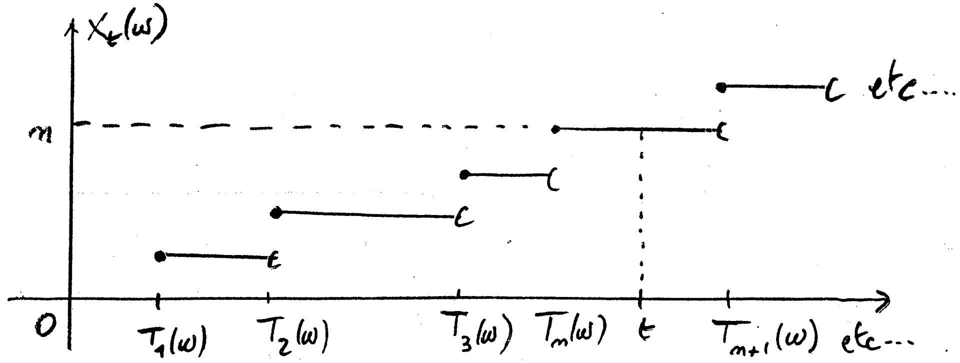
<u>Construction d'un processus de Poisson à partir d'une suite $(\tau_i)_{i\geq 0}$ de v.a. i.i.d de loi exponentielle de paramètre $\lambda > 0$</u> :

Etant donné une suite $(\tau_i)_{i\geq 0}$ de v.a. i.i.d de loi $\mathcal{E}(\lambda)$, $\lambda > 0$ considérons les v.a. $T_0 = 0$ et

$$T_1 = \tau_1$$
$$T_2 = \tau_1 + \tau_2$$
$$\vdots$$
$$T_n = \tau_1 + \tau_2 + \dots + \tau_n$$
$$\vdots$$

Pour tout $t > 0$, considérons la v.a. X_t défini par:

$X_t = n$ si $T_n \le t < T_{n+1}$ $(n \in \mathbb{N})$. Pour $\omega \in \Omega$ fixé, la fonction $t \mapsto X_t(\omega)$ a le graphique suivant :



$t \mapsto X_t(\omega)$ est une fonction étagée croissante dont les sauts d'amplitude 1, ont lieu aux instants $T_n(\omega)$ $(n \in \mathbb{N})$. On montre (voir les TD) que la famille $(X_t)_{t \ge 0}$ est un processus de Poisson d'intensité $\lambda > 0$.

Propriété de Markov et probabilités de transition d'un processus de Poisson.

Considérons un processus de Poisson $(X_t)_{t\ge 0}$ d'arrivées purement aléatoires d'intensité $\lambda > 0$ comme décrit page 6. L'hypothèse H_2 signifie que quelle que soit la suite finie d'instants $0 \le t_1 < t_2 < \dots < t_n$, les v.a. $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ sont indépen-
-dants. On dit que le processus (X_t) est à

accroissements indépendants

Théorème : pour toute suite finie d'instants $0 \le t_1 < t_2 < \dots < t_n$ et toute suite finie d'entiers $k_1 \le k_2 \le \dots \le k_n$, on a

$$\mathbb{P}(X_{t_n} = k_n \mid X_{t_{n-1}} = k_{n-1}, X_{t_{n-2}} = k_{n-2}, \dots, X_{t_1} = k_1)$$
$$= \mathbb{P}(X_{t_n} = k_n \mid X_{t_{n-1}} = k_{n-1})$$

(propriété de Markov du processus de Poisson)

démonstration : Par définition de la probabilité conditionnelle, on a :

$$\mathbb{P}(X_{t_n} = k_n \mid X_{t_{n-1}} = k_{n-1}, \dots, X_{t_1} = k_1) =$$
$$\frac{\mathbb{P}(X_{t_n} = k_n, X_{t_{n-1}} = k_{n-1}, \dots, X_{t_1} = k_1)}{\mathbb{P}(X_{t_{n-1}} = k_{n-1}, \dots, X_{t_1} = k_1)}$$
$$= \frac{\mathbb{P}(X_{t_1} = k_1, X_{t_2} - X_{t_1} = k_2 - k_1, \dots, X_{t_n} - X_{t_{n-1}} = k_n - k_{n-1})}{\mathbb{P}(X_{t_1} = k_1, X_{t_2} - X_{t_1} = k_2 - k_1, \dots, X_{t_{n-1}} - X_{t_{n-2}} = k_{n-1} - k_{n-2})}$$
$$= \frac{\mathbb{P}(X_{t_1} = k_1)\mathbb{P}(X_{t_2} - X_{t_1} = k_2 - k_1) \dots \mathbb{P}(X_{t_n} - X_{t_{n-1}} = k_n - k_{n-1})}{\mathbb{P}(X_{t_1} = k_1)\mathbb{P}(X_{t_2} - X_{t_1} = k_2 - k_1) \dots \mathbb{P}(X_{t_{n-1}} - X_{t_{n-2}} = k_{n-1} - k_{n-2})}$$
$$= \mathbb{P}(X_{t_n} - X_{t_{n-1}} = k_n - k_{n-1}) \quad \circledast$$
$$= \mathbb{P}(X_{t_n} = k_n \mid X_{t_{n-1}} = k_{n-1}) \quad \text{cqfd.}$$

Définition : pour $s \le t$ et des entiers $i \le j$,

$\mathbb{P}(X_t = j \mid X_s = i)$

est la probabilité de transition partant de i à l'instant s, d'aller en j à l'instant t

Homogénéité de la probabilité de transition

D'après ce qu'on a trouvé en ⊛ :

$$\mathbb{P}(X_t = j \mid X_s = i) = \mathbb{P}(X_t - X_s = j - i)$$

$$= P_{j-i}(t-s) \left(= e^{-\lambda(t-s)} \frac{(\lambda(t-s))^{j-i}}{(j-i)!} \right)$$

= la probabilité d'avoir $j-i$ arrivées pendant un intervalle de temps de longueur $t-s$ d'après l'hypothèse H_1. Le processus de Markov $(X_t)_{t\ge 0}$ est donc homogène dans le temps

Notons qu'il est aussi homogène en espace (c'est la propriété d'accroissements indépendants)

III Processus de naissance et mort

On considère ici un système d'attente où on a un mécanisme d'arrivées (naissances) et un mécanisme de sorties (morts). On s'intéresse à la longueur X_t de la ligne d'attente à l'instant t (i.e. au nombre de clients présents dans le système au temps t). On va se

restreindre à des considérations élémentaires.

Ici $(X_t)_{t\ge 0}$ sera un processus de Markov à espace d'états $\mathbb{N} = \{0, 1, 2, \dots\}$. Si à l'instant t le processus est dans l'état n, il peut après un temps d'attente aléatoire, passer en l'un des états voisins $n+1$ (naissance) ou $n-1$ (décès) ainsi $(X_t)_{t\ge 0}$ est l'analogue en temps continu d'une marche aléatoire sur $\mathbb{N}$. On dira que $(X_t)_{t\ge 0}$ est un processus de naissance et mort s'il vérifie les conditions suivantes :

1) $(X_t)_{t\ge 0}$ est un processus de Markov homogène sur $\mathbb{N}$ (i.e. ses probabilités de transition

$$P_{ij}(t) = \mathbb{P}(X(t+s) = j \mid X(s) = i)$$

sont homogènes en temps)

2) Il existe deux suites de nombres positifs $(\lambda_n)_{n\ge 0}$ et $(\mu_n)_{n\ge 0}$ telles que $P_{ij}(t)$ vérifie les conditions suivantes :

2a $\quad P_{i\,i+1}(h) = \lambda_i h + o(h)$ si $i \ge 0$

2b $\quad P_{i\,i-1}(h) = \mu_i h + o(h)$ si $i \ge 1$

2c $\quad P_{ii}(h) = 1 - (\lambda_i + \mu_i) h + o(h)$ si $i \ge 0$

2d $\quad P_{ij}(0) = \delta_{ij}$

2e $\mu_0 = 0$, $\lambda_0 > 0$, $\mu_n, \lambda_n > 0$ si $n > 0$.

λ_n (resp. μ_n) est appelé intensité ou taux de naissance (resp. de décès) en l'état n.