

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

DÉPARTEMENT DE MATHÉMATIQUES



Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : Statistique

Par

HABCHI Oumaima

Titre :

Estimation de l'intensité d'un processus de  
Poisson homogène

Membres du Comité d'Examen :

Dr. ABDELLI Jihane	UMKB	Président
Dr. DJABER Ibtissem	UMKB	Encadreur
Dr. SOLTANE Louiza	UMKB	Examinatrice

Juin 2021

## Dédicace

*À mes parents,*

*À mes soeurs et frères,*

*À tous ceux qui me sont proches.*

## REMERCIEMENTS

*Je tiens tout d'abord à exprimer ma profonde reconnaissance pour mon encadreur  
Madame Djaber Ibtissem qui a dirigé l'ensemble de ce travail. Je la remercie  
profondément pour ses remarques toujours pertinentes ainsi que sa gentillesse  
omniprésente.*

*Mes sincères remerciements à mes professeurs de mathématiques de l'Université de  
Mohamed Khider de Biskra.*

# Table des matières

<b>Remerciements</b>	iii
<b>Table des matières</b>	iv
<b>Table des figures</b>	v
<b>Introduction</b>	1
<b>1 Notions sur les processus stochastiques</b>	4
1.1 Loi exponentielle et loi de Poisson . . . . .	4
1.2 Processus stochastiques . . . . .	7
1.3 Processus de comptage . . . . .	10
<b>2 Estimation de l'intensité d'un processus de Poisson homogène</b>	13
2.1 Processus de Poisson homogène . . . . .	13
2.1.1 Distribution conditionnelle des instants d'arrivée . . . . .	23
2.2 Estimation par la méthode du maximum de vraisemblance . . . . .	26
2.3 Estimation de l'intensité d'un processus de Poisson homogène . . . . .	28
2.3.1 Comportement asymptotique . . . . .	30
<b>Conclusion</b>	34

<b>Bibliographie</b>	<b>35</b>
<b>Annexe A : Logiciel <i>R</i></b>	<b>36</b>
<b>Annexe B : Abréviations et Notations</b>	<b>38</b>

# Table des figures

1.1 Probabilite d'une (v.a) de loi de Poisson de paramètre 0.75,3, 6 et 10	5
1.2 Densité de la loi exponentielle de paramètres 0.5,1,1.5et2 . . . . .	6
1.3 Processus de comptage . . . . .	12
2.1 Trajectoires d'un processus de Poisson d'intensité $\lambda = 3, 7$ . . . . .	14

# Introduction

**E**n probabilité et statistiques, un processus de Poisson est un type d'objet aléatoire connu sous le nom de processus ponctuel qui se compose de points positionnés de manière aléatoire situés sur un espace mathématique sous-jacent. Le processus a des propriétés mathématiques pratiques, ce qui l'a conduit à être fréquemment défini dans l'espace euclidien et utilisé comme modèle mathématique pour des processus apparemment aléatoires dans de nombreuses disciplines telles que l'astronomie, la biologie, l'écologie, géologie, physique, traitement d'images et télécommunications.

Le processus de Poisson est souvent défini sur la ligne réelle jouant un rôle important dans le domaine de la théorie des files d'attente où il est utilisé pour modéliser certains événements aléatoires se produisant dans le temps tels que l'arrivée de clients dans un magasin ou des appels téléphoniques dans un magasin.

Dans le plan, le processus ponctuel, également connu sous le nom de processus spatial de Poisson, peut représenter des objets dispersés tels que des utilisateurs dans un réseau sans fil.

Le processus ponctuel de Poisson a la propriété que chaque point est stochastiquement indépendant de tous les autres points du processus, c'est pourquoi il est parfois appelé processus purement ou complètement aléatoire.

Le processus est nommé d'après le mathématicien français Simeon Denis Poisson en raison du fait que si une collection de points aléatoires dans un espace forme un

processus de Poisson, alors le nombre de points dans une région de taille finie est directement lié à la distribution de Poisson, mais Poisson, cependant, n'a pas étudié le processus, qui est apparu indépendamment dans plusieurs contextes différents.

Le processus est défini avec un objet unique, qui, selon le contexte, peut être une constante, une fonction intégrable ou, plus généralement, une mesure de Radon. Si cet objet est une constante, alors le processus résultant est appelé un processus de Poisson ( homogène ou stationnaire ). Sinon, le paramètre dépend de sa localisation dans l'espace sous-jacent, ce qui conduit au processus de Poisson inhomogène ou non homogène.

Afin d'utiliser un processus de Poisson dans les applications, une étape clé consiste à estimer sa fonction d'intensité à partir d'une séquence donnée d'événements observés. L'estimation de la fonction d'intensité d'un processus de Poisson a été largement étudiée et diverses méthodes d'estimation ont été proposées.

Si l'intensité peut être supposée avoir une forme paramétrique connue, alors des méthodes basées sur la vraisemblance peuvent être utilisées pour estimer les paramètres du modèle.

Cependant, dans de nombreux cas, la forme de l'intensité est inconnue et l'estimation nécessite la mise en œuvre de méthodes non paramétriques. Les méthodes d'estimation non paramétriques offrent plus de flexibilité que les méthodes paramétriques et peuvent mieux caractériser la fonction d'intensité sous-jacente. Dans le cas lorsque les connaissances préalables sur le processus ou la forme de l'intensité sont connues, les méthodes bayésiennes peuvent être adoptées et elles conduisent souvent à une estimation plus précise.

L'objectif de ce mémoire est de donner une présentation de processus de Poisson et ces propriétés fondamentales, estimer l'intensité  $\lambda$  de ce processus par la méthode de maximum de vraisemblance et étudier le comportement asymptotique de cet estimateur. De plus on utilise le logiciel statistique R pour définir la notion des trajectoires



de ce processus.

Ce mémoire est divisé en deux chapitres

Dans le premier chapitre, on commence par rappeler la loi exponentielle et la loi de Poisson, en suite on présente des notions générales sur les processus stochastiques et leurs propriétés, puis on définit les processus de comptages (car le processus de Poisson est le plus simple processus de comptage utilisé dans la modélisation des files d'attente).

Le deuxième chapitre est consacré à définir le processus de Poisson homogène et leurs propriétés principales, en suite une rappelle sur la méthode de maximum de vraisemblance. En fin on utilise cette méthode pour estimer l'intensité  $\lambda$  de ce processus et étudier le comportement asymptotique de cet estimateur

# Chapitre 1

## Notions sur les processus stochastiques

L'objectif de ce chapitre est de rappeler des définitions et présenter quelques propriétés sur la loi exponentielle, la loi de Poisson, les processus stochastiques, et le processus de comptage.

### 1.1 Loi exponentielle et loi de Poisson

**Définition 1.1.1** Une variable aléatoire  $X$  est dite suivre une loi de Poisson de paramètre  $\lambda > 0$ , et on note  $X \sim \mathcal{P}(\lambda)$ , si

$$P(X = k) = \frac{\lambda^k}{k!} \exp(-\lambda), \quad k \in \mathbb{N}.$$

$X$  admet alors une espérance et une variance

$$E[X] = \lambda \quad \text{et} \quad \text{Var}[X] = \lambda.$$

La figure (1.1) représente la probabilité associée à la loi de Poisson des paramètres 0.75,3,6,10, ils ont été obtenus dans le logiciel R par la commande `dpois()`.

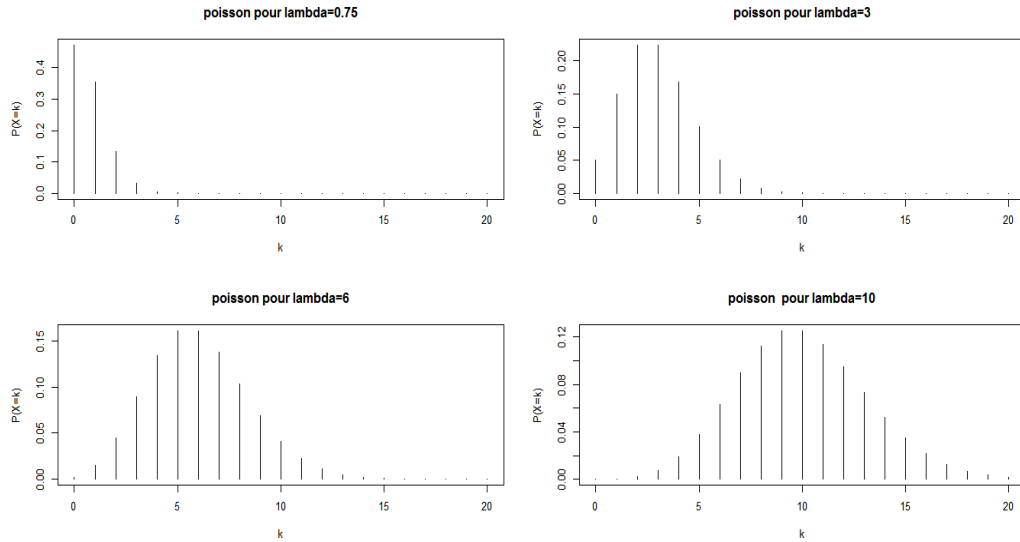


FIG. 1.1 – Probabilité d’une (v.a) de loi de Poisson de paramètre 0.75,3, 6 et 10

**Définition 1.1.2** *une variable aléatoire  $X$  absolument continue suit une loi exponentielle de paramètre  $\lambda > 0$ , et on note  $X \sim \mathcal{E}(\lambda)$ , si la densité de probabilité est*

$$f(x) = \begin{cases} \lambda \exp(-\lambda x); & \text{si } x \geq 0 \\ 0; & \text{si } x < 0 \end{cases}$$

*et sa fonction de répartition est donnée par*

$$F_{\mathbf{X}}(t) = \begin{cases} 1 - \exp(-\lambda t); & \text{si } t \geq 0 \\ 0; & \text{si } t < 0 \end{cases}$$

*$X$  admet alors une espérance et une variance*

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

On peut représenter des exemples de densité associées à la loi exponentielle de paramètres différentes, pour obtenir la densité de  $X$  par le langage R on utilise la commande `dexp()`.

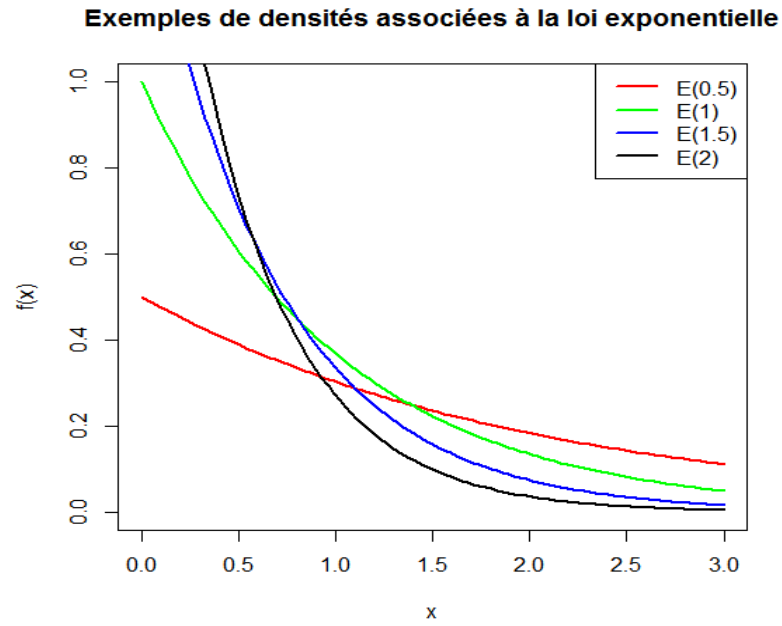


FIG. 1.2 – Densité de la loi exponentielle de paramètres 0.5,1,1.5et2

**Propriété 1.1.1** (*Absence de mémoire de la loi exponentielle*)

*Si la variable aléatoire  $X$  suit une loi exponentielle alors*

$$P(X > s + t \mid X > s) = P(X > t), \text{ pour tous } s, t \geq 0.$$

**Preuve.** Supposons tout d'abord que  $X$  suit une loi exponentielle de paramètre  $\lambda$ .

On a

$$\begin{aligned}
 P(X > s + t \mid X > s) &= \frac{P(X > s + t, X > s)}{P(X > s)} \\
 &= \frac{P(X > s + t)}{P(X > s)} \\
 &= \frac{1 - P(X \leq s + t)}{1 - P(X \leq s)} \\
 &= \frac{\exp(-\lambda(s + t))}{\exp(-\lambda s)} \\
 &= \exp(-\lambda t) = P(X > t).
 \end{aligned}$$

■

## 1.2 Processus stochastiques

On s'intéresse à l'évolution d'un phénomène, d'une variable au cours du temps. Pour modéliser ces phénomènes on utilise le processus stochastique.

Soit  $(\Omega, \mathcal{F}, P)$  un espace de probabilité, on désigne par  $\mathbb{T}$  un ensemble arbitraire s'appelle l'ensemble des indices; souvent on prend  $\mathbb{T} = \mathbb{R}, \mathbb{R}_+, \mathbb{R}^N$  (où  $N \in \mathbb{N}^*$ ),  $\mathbb{R}_+^N, \mathbb{N}, \mathbb{Z}$ , ou encore une partie de l'un de ces ensembles, et on désigne par  $\mathbb{E}$  un espace métrique muni de la  $\sigma$ -algèbre borélienne notée par  $\mathfrak{B}(\mathbb{E})$  cet espace métrique est appelé l'espace d'états, souvent on prend  $\mathbb{E} = \mathbb{R}, \mathbb{C}, \mathbb{R}^d$  (où  $d \in \mathbb{N}^*$ ),  $\mathbb{C}^d$ , un ensemble fini ou un ensemble infini dénombrable. Donc l'espace  $(\mathbb{E}, \mathfrak{B}(\mathbb{E}))$  est un espace mesurable.

**Définition 1.2.1** *Un processus aléatoire (stochastique) est une famille de variables aléatoires  $(X_t)_{t \in \mathbb{T}}$  indexées par l'ensemble  $\mathbb{T}$  des temps dénombrable ou continue, définies sur un espace de probabilité  $(\Omega, \mathcal{F}, P)$ , et à valeurs dans  $(\mathbb{E}, \mathfrak{B}(\mathbb{E}))$ .*

**Remarque 1.2.1** Lorsque  $\mathbb{T} = \mathbb{N}$  ou une partie de  $\mathbb{N}$  on dit qu'on a un processus à temps discret ou une suite aléatoire, et si  $\mathbb{T} = \mathbb{R}$  ou une partie de  $\mathbb{R}$  on dit qu'on a un processus à temps continue.

**Remarque 1.2.2** Le processus  $(X_t)_{t \in \mathbb{T}}$  dépend de deux paramètres le temps  $t \in \mathbb{T}$  et de l'aléatoire  $\omega \in \Omega$ .

- Pour tout  $t$  fixé  $\omega \in \Omega \longrightarrow X_t(\omega)$  est une variable aléatoire sur l'espace de probabilité  $(\Omega, \mathcal{F}, P)$ .
- Pour  $\omega$  est fixé,  $t \in \mathbb{T} \longrightarrow X_t(\omega)$  est une fonction à valeurs réelles, appelée trajectoire du processus.

**Définition 1.2.2** Un processus aléatoire  $(X_t)_{t \in \mathbb{T}}$  est dit "Càd-làg" si ses trajectoires ( $t \in \mathbb{T} \longrightarrow X_t(\omega)$ ) sont continue à droite et admettent en tout point, une limite à gauche.

**Définition 1.2.3** (Lois finie dimensionnelles)

Les lois finie dimensionnelles d'un processus stochastique  $(X_t)_{t \in \mathbb{T}}$  sont les lois de tous les vecteurs aléatoires  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ ,  $n \in \mathbb{N}^*$ , pour  $t_1, \dots, t_n \in \mathbb{T}$ , donnée par

$$F_{(X_{t_1}, X_{t_2}, \dots, X_{t_n})}(x_1, \dots, x_n) = P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n), \quad t_1, \dots, t_n \in \mathbb{T}.$$

**Remarque 1.2.3** Ces lois caractérisent le processus stochastique.

**Définition 1.2.4** Deux processus  $X$  et  $Y$  ont même lois s'ils ont même lois fini-dimensionnelles : pour tout  $t_1, \dots, t_n \in T$  et  $n \in \mathbb{N}^*$

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{=} (Y_{t_1}, \dots, Y_{t_n}).$$

**Définition 1.2.5** La variable aléatoire  $(X_{t_i} - X_{t_j})$  où  $t_i < t_j$  est l'accroissement du processus sur l'intervalle  $[t_i, t_j[$ .

Un processus aléatoire  $(X_t)_{t \in \mathbb{T}}$  à valeurs dans  $\mathbb{R}$  est un processus à accroissement indépendants si  $n \in \mathbb{N}^*$ , pour tout  $0 < t_1 < t_2 < \dots < t_n$ , les variables aléatoires  $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$  sont deux à deux indépendantes.

**Définition 1.2.6** On dit que  $(X_t)_{t \in \mathbb{T}}$  est un processus à accroissements stationnaires si pour tout  $h \geq 0$

$$X_{t+h} - X_t \stackrel{\mathcal{L}}{=} X_h.$$

i.e la loi des accroissements ne dépend pas de  $t$ .

**Définition 1.2.7** On dit que  $(X_t)_{t \in \mathbb{T}}$  est un processus strictement stationnaire si pour tout  $h > 0$  et tout  $t_1, \dots, t_n \geq 0$ , on a

$$(X_{t_1+h}, \dots, X_{t_n+h}) \stackrel{\mathcal{L}}{=} (X_{t_1}, \dots, X_{t_n}).$$

**Définition 1.2.8** Un processus  $(X_t)_{t \in \mathbb{T}}$  est continue en probabilité au point  $t$ , si  $\forall \varepsilon > 0$

$$P(\omega : |X_{t+h} - X_t| > \varepsilon) \longrightarrow 0 \quad \text{si } h \longrightarrow 0.$$

**Définition 1.2.9** Un processus  $(X_t)_{t \in \mathbb{T}}$  est dit localement continue en probabilité si pour  $t \geq 0$ , on a

$$\lim_{h \rightarrow 0} P(X_{t+h} - X_t \geq 1) = 0.$$

**Propriétés 1.2.1** Soit  $(X_t) \in L_2$ . Dans ce cas on peut définir les fonctions suivantes

○ Fonction moyenne d'un processus aléatoire  $X_t$  est

$$m_X(t) = E[X_t].$$

○ Variance d'un processus aléatoire  $X_t$  est

$$\text{Var}[X_t] = \sigma_t^2.$$

◦ *Fonction de corrélation*

$$R_X(s, t) = \text{Cov}(X_s, X_t) = E[(X_s - m_X(s))(X_t - m_X(t))].$$

◦ *Fonction d'autocorrélation*

$$\text{corr}(X_s, X_t) = \frac{R_X(s, t)}{\sqrt{\text{Var}[X_s]\text{Var}[X_t]}} = \frac{\text{Cov}(X_s, X_t)}{\sqrt{\text{Var}[X_s]\text{Var}[X_t]}}.$$

**Remarque 1.2.4** *Si  $\forall t \in \mathbb{T}$  la (v.a)  $X_t$  est intégrable et  $E[X_t] = 0$ , alors le processus  $(X_t)$  est centré.*

### 1.3 Processus de comptage

On peut étudier un phénomène aléatoire survenu au cours de temps par un processus stochastique comme le processus de comptage  $\{N(t), t \geq 0\}$ , et ces phénomènes sont du genre :

- Tremblement de terre,
- Accident de voiture,
- arrivées des clients dans un magasin.

**Définition 1.3.1** *Soit  $\{N(t), t \geq 0\}$  le nombre d'événements survenus dans l'intervalle  $[0, t]$ . Le processus stochastique  $\{N(t), t \geq 0\}$  est appelé processus de comptage, par exemple  $N(t)$  représente le nombre de clients arrivés devant un guichet au cours de la période  $[0, t]$ .*



Le processus de comptage  $N(t)$  possède les propriétés suivantes

- Propriétés 1.3.1**
1.  $N(0) = 0$  et  $N(t) \in \mathbb{N}$ ,
  2. Le processus de comptage  $N(t)$  est à valeurs entières ( $N(t) < \infty$ ) et positives,
  3. La fonction  $N(t)$  est croissante :  $N(t_2) - N(t_1) \geq 0$  si  $0 \leq t_1 < t_2$ . de plus  $N(t_2) - N(t_1)$  représente le nombre d'événements survenus dans l'intervalle  $[t_1, t_2[$ ,
  4. Les trajectoires  $t \rightarrow N_t(\omega)$  sont continues à droite et admettent une limite à gauche.

**Propriétés 1.3.2** Un processus de comptage  $\{N(t), t \geq 0\}$  est dit à accroissements stationnaires si la loi de probabilité du nombre d'événements se produisant dans un intervalle de temps donné ne dépend que de la longueur de cet intervalle.

On associe au processus de comptage  $N(t)$  ses deux processus des temps successifs  $\{T_n, n \in \mathbb{N}\}$  et  $\{S_n, n \in \mathbb{N}\}$  tel que

- o Le processus  $\{T_n, n \in \mathbb{N}\}$ , appelé processus des temps d'occurrence (d'arrivées) est continue à temps discret définie par

$$T_0 := 0 \quad \text{et} \quad T_n := S_1 + \dots + S_n (n \geq 1).$$

Par conséquent, la variable aléatoire  $T_n$  ( $n \geq 1$ ) est le temps d'occurrences du  $n^{ieme}$  événement (à partir de  $t = 0$ ).

- o Le processus  $\{S_n, n \in \mathbb{N}\}$ , appelé processus des temps d'inter\_occurrence (d'inter\_arrivées), la variable aléatoire  $S_n$  ( $\forall n \in \mathbb{N}$ ) est le temps d'attente entre les  $(n - 1)^ieme$  et  $(n)^ieme$  occurrences

$$S_n := T_n - T_{n-1}.$$

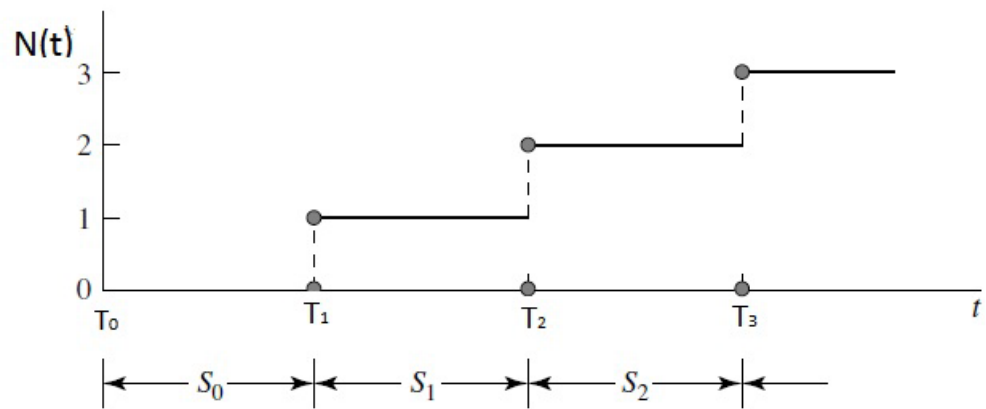


FIG. 1.3 – Processus de comptage

**Remarque 1.3.1** *On remarque que le nombre d'événements  $N(t)$  dépend des  $T_n$  par les relations suivantes*

- $\{N_t = n\} = \{T_n \leq t < T_{n+1}\}$ ,
- $\{N_t \geq n\} = \{T_n \leq t\}$ .

# Chapitre 2

## Estimation de l'intensité d'un processus de Poisson homogène

L'objectif de ce chapitre est de présenter le processus de Poisson homogène ainsi que ces propriétés fondamentales, et estimer l'intensité par la méthode du maximum de vraisemblance.

### 2.1 Processus de Poisson homogène

Le processus de Poisson est un processus à temps continue et à valeurs entières positives, on dit encore que c'est un processus de comptage utilisé pour modéliser des files d'attente par exemple : les appels dans un centrale téléphonique, les arrivées de clients devant un guichet, pannes de machines dans une usine.

**Définition 2.1.1** *Un processus de comptage  $\{N(t), t \geq 0\}$  est appelé processus de Poisson d'intensité  $\lambda > 0$ , s'il vérifie les propriétés suivantes*

1.  $N(0) = 0$ ,
2. *Le processus est à accroissements indépendants,*

3. Le nombre d'occurrences se produisant dans un intervalle de temps de longueur  $t \geq 0$  suit la loi de Poisson de paramètre  $\lambda t$ , c'est-à-dire,  $\forall s, t \geq 0$ , on a

$$P(N(s+t) - N(s) = n) = \exp(-\lambda t) \frac{(\lambda t)^n}{n!} \quad (n \geq 0).$$

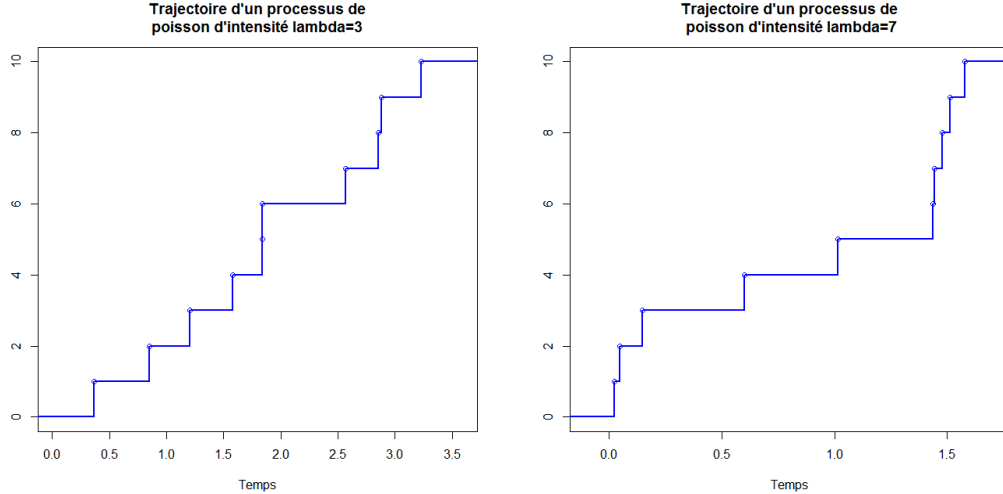


FIG. 2.1 – Trajectoires d'un processus de Poisson d'intensité  $\lambda = 3, 7$

**Propriété 2.1.1** *Un processus de Poisson est à accroissements stationnaires.*

**Preuve.** Pour deux intervalles de temps  $]a, c], ]b, d]$  avec  $a < b \leq c < d$ , et comme le processus est à accroissements indépendants, on a

$$\begin{aligned} & P(N(c) - N(a) = k, N(d) - N(b) = l) = \\ &= \sum P(N(b) - N(a) = i_1, N(c) - N(b) = i_2, N(d) - N(c) = i_3) \\ &= \sum P(N(b) - N(a) = i_1) P(N(c) - N(b) = i_2) P(N(d) - N(c) = i_3) \\ &= \sum \exp(-\lambda(d-a)) \frac{(\lambda(b-a))^{i_1}}{i_1!} \frac{(\lambda(c-b))^{i_2}}{i_2!} \frac{(\lambda(d-c))^{i_3}}{i_3!}. \end{aligned}$$

où la somme est sur toutes les suites  $(i_1, i_2, i_3)$  d'entier tel que  $i_1 + i_2 = k$  et  $i_2 + i_3 = l$ , or cette expression ne dépend que des longueurs des intervalles de temps. ■

**Propriété 2.1.2** Soit  $\{N(t), t \geq 0\}$  un processus de Poisson d'intensité  $\lambda > 0$ .

Alors, lorsque  $h$  tend vers 0

i.  $P(N(h) = 1) = \lambda h + o(h)$ ,

ii.  $P(N(h) \geq 2) = o(h)$ .

**Preuve.**

i. On a

$$P(N(h) = 1) = \lambda h \exp(-\lambda h).$$

On fait le développement limité de

$$\exp(-\lambda h) \text{ pour } h \rightarrow 0,$$

alors

$$P(N(h) = 1) = \lambda h(1 + o(h)) = \lambda h + o(h).$$

ii. En effet

$$P(N(h) = 0) = \exp(-\lambda h).$$

On fait le développement limité de

$$\exp(-\lambda h) \text{ pour } h \rightarrow 0,$$

on obtient

$$P(N(h) = 0) = 1 - \lambda h + o(h),$$

alors

$$\begin{aligned}
 P(N(h) \geq 2) &= 1 - P(N(h) < 2) \\
 &= 1 - [P(N(h) = 0) + P(N(h) = 1)] \\
 &= o(h), \quad \text{lorsque } h \rightarrow 0.
 \end{aligned}$$

■

**Exemple 2.1.1** On considère un processus de Poisson  $\{N(t), t \geq 0\}$  d'intensité  $\lambda = 2$ , la variable aléatoire  $N(t)$  représente par exemple le nombre d'appels à un centrale téléphonique entre les dates 0 et  $t$ , au bout de 1 unité de temps, le nombre d'appels est  $N(1)$  et suit donc une loi de Poisson de paramètre 2, au bout de 10 unités de temps, le nombre d'appels est  $N(10)$  et suit donc une loi de Poisson de paramètre 20.

**Définition 2.1.2** Soit  $T_n (n \geq 1)$  un temps d'arrivée du  $n^{\text{ième}}$  évènements définie par

$$T_n = S_1 + \dots + S_n \quad (n \geq 1),$$

où  $(S_n)_{n \geq 1}$  est une suite de variables aléatoires indépendantes s'appelle processus des inter-arrivées définie par

$$S_n = T_n - T_{n-1} (n \geq 1).$$

**Théorème 2.1.1** Soit

$$T_n = \inf\{t \geq 0, N(t) \geq n\} \quad \text{et} \quad S_n = T_n - T_{n-1} (n \geq 1).$$

Alors

1.  $(S_n)_{n \geq 1}$  est une suite des (v.a) i.i.d de loi  $\mathcal{E}(\lambda)$ ,  $\lambda > 0$ , et sa densité donnée par

$$f_{(S_1, \dots, S_n)}(s_1, \dots, s_n) = \lambda^n \exp(-\lambda s_n) \mathbf{1}_{0 < s_1 < \dots < s_n}.$$

2.  $T_n = S_1 + \dots + S_n$  suit la loi Gamma  $\Gamma(n, \lambda)$  de densité

$$f_{T_n}(t) = \begin{cases} \frac{\lambda}{(n-1)!} (\exp(-\lambda t)) (\lambda t)^{n-1} & \text{si } t \geq 0 \\ 0 & \text{si non.} \end{cases}$$

**Preuve.**

1. Soit  $n \in \mathbb{N}^*$

**Étape 1 : Changement de variable :** supposons que le vecteur aléatoire  $(T_1, \dots, T_n)$  soit à densité, de densité  $\varphi$ . Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  continue bornée, alors comme  $T_1 \leq T_2 \leq \dots \leq T_n$  par un changement de variable  $t_k = s_1 + \dots + s_k$  ( $1 \leq k \leq n$ ) de jacobien 1 (la matrice jacobienne est triangulaire), on a

$$\begin{aligned} E[f(S_1, \dots, S_n)] &= E\left[\mathbf{1}_{0 < T_1 \leq T_2 \leq \dots \leq T_n} f(T_1, \dots, T_n - T_{n-1})\right] \\ &= \int_{0 < t_1 \leq t_2 \leq \dots \leq t_n} f(t_1, \dots, t_n - t_{n-1}) \varphi(t_1, \dots, t_n) dt_1, \dots, dt_n \\ &= \int_{0 < s_1 \leq s_2 \leq \dots \leq s_n} f(s_1, \dots, s_n) \varphi(s_1, \dots, s_1 + \dots + s_n) ds_1 \dots ds_n. \end{aligned}$$

Donc

$$\psi : (s_1, \dots, s_n) \rightarrow \varphi(s_1, s_1 + s_2, \dots, s_1 + \dots + s_n);$$

est la densité de  $(S_1, \dots, S_n)$ .

**Étape 2 : calcul de la densité de  $(T_1, \dots, T_n)$**

Soit l'évènement

$$A_n = \{T_1 \in [t_1, t_1 + h_1[, T_2 \in [t_2, t_2 + h_2[, \dots, T_n \in [t_n, t_n + h_n]\},$$

où  $0 < t_1 < t_1 + h_1 < t_2 < t_2 + h_2 < \dots < t_n < t_n + h_n$ , alors  $A_n$  est la réunion des évènements

-zéro top dans  $[0, t_1[$  et exactement un top dans  $[t_1, t_1 + h_1[$

-zéro top dans  $[t_1 + h_1, t_2[$  et exactement un top dans  $[t_2, t_2 + h_2[$

⋮

-zéro top dans  $[t_{n-1} + h_{n-1}, t_n[$  et au moins un top dans  $[t_n, t_n + h_n[$ .

Or, le processus étant à accroissements indépendants, les variables aléatoires «nombre de tops» dans des intervalles disjoints sont indépendantes de sorte que

$$\begin{aligned}
 P(A_n) &= P(t_1 < T_1 < t_1 + h_1, \dots, t_n < T_n < t_n + h_n) \\
 &= P(N_{t_1} = 0, N_{t_1+h_1} - N_{t_1} = 1, \dots \\
 &\quad \dots, N_{t_n} - N_{t_{n-1}+h_{n-1}} = 0, N_{t_n+h_n} - N_{t_n} \geq 1) \\
 &= P(N_{t_1} = 0) \times P(N_{t_1+h_1} - N_{t_1} = 1) \times P(N_{t_2} - N_{t_1+h_1} = 0) \\
 &\quad \times P(N_{t_2+h_2} - N_{t_2} = 1) \times \dots \times P(N_{t'_n} - N_{t_{n-1}+h_{n-1}} = 0) \\
 &\quad \times P(N_{t_n+h_n} - N_{t_n} \geq 1) \\
 &= \exp(-\lambda t_1) \exp(-\lambda h_1) \lambda h_1 \exp(-\lambda(t_2 - (t_1 + h_1))) \exp(-\lambda h_2) \lambda h_2 \dots \\
 &\quad \dots \exp(-\lambda(t_n - (t_{n-1} + h_{n-1}))) (1 - \exp(-\lambda h_n)) \\
 &= \exp(-\lambda t_n) (1 - \exp(-\lambda h_n)) \lambda^{n-1} h_1 \dots h_{n-1}.
 \end{aligned}$$

Pour conclure, il suffit de remarquer que

$$P(A_n) = \int_{\xi=t_0}^{t_1+h_1} \dots \int_{\xi=t_n}^{t_n+h_n} 1_{0 < \xi_1 \leq \dots \leq \xi_n} \lambda^n \exp(-\lambda \xi_n) d\xi_1 \dots d\xi_n.$$

Ceci valant pour tous les pavés  $[t_1, t_1 + h_1[ \times \dots \times [t'_n, t_n + h_n[$ , qui constituent une classe stable par intersection engendrant  $\mathfrak{B}(\mathbb{R}^n)$  donc  $(T_1, \dots, T_n)$  a pour densité

$$1_{0 < \xi_1 \leq \dots \leq \xi_n} \lambda^n \exp(-\lambda \xi_n).$$



**Conclusion :** Selon la première étape, la densité de  $(S_1, \dots, S_n)$  est

$$(s_1, \dots, s_n) \longrightarrow \lambda^n \exp(-\lambda s_1) \dots \exp(-\lambda s_n) \mathbf{1}_{\mathbb{R}^n}(s_1, \dots, s_n).$$

En calculant les densités marginales, on constate immédiatement que

$$f_{(S_1, \dots, S_n)}(s_1, \dots, s_n) = f_{S_1}(s_1) \dots f_{S_n}(s_n);$$

en d'autres termes,  $S_1, \dots, S_n$  sont indépendantes.

2. On peut démontrer que si  $S_1, \dots, S_n$  sont des (v.a) indépendantes de loi  $\mathcal{E}(\lambda)$ , alors

$$T_n = S_1 + \dots + S_n,$$

suit la loi Gamma  $\Gamma(n, \lambda)$ .

**Par récurrence**

Pour  $n = 1$ , la fonction de densité de  $T_1$  (où  $T_1 = S_1$ ) est

$$\begin{aligned} f_{T_1}(t) &= \lambda \exp(-\lambda t) \\ &= \frac{\lambda}{0!} \exp(-\lambda t) (\lambda t)^0, \end{aligned}$$

la propriété est vraie pour  $n = 1$ ,

On suppose alors que pour une valeur de  $n$  la propriété est vraie c'est-à-dire, la fonction de densité de  $T_n$  (où  $T_n = S_1 + \dots + S_n$ ) est

$$f_{T_n}(t) = \begin{cases} \frac{\lambda}{(n-1)!} (\exp(-\lambda t)) (\lambda t)^{n-1}, & \text{si } t \geq 0 \\ 0, & \text{si non.} \end{cases}$$

Il reste à démontrer que la propriété est vraie pour  $n + 1$ .  $\forall t \geq 0$ , on a

$$T_{n+1} = S_1 + S_2 + \dots + S_n + S_{n+1} = T_n + S_{n+1}.$$

Si  $T_n$  admet comme fonction de densité  $f_{T_n}$  et  $S_{n+1}$  admet comme fonction de densité  $f_{S_{n+1}}$ , avec  $T_n$  et  $S_{n+1}$  sont indépendantes, la fonction de densité de  $T_n + S_{n+1}$  est **le produit de convolution**, alors

$$\begin{aligned} f_{T_{n+1}}(t) &= \{f_{T_n} \star f_{S_{n+1}}\}(t) \\ &= \int_{-\infty}^{+\infty} f_{T_n}(u) f_{S_{n+1}}(t-u) du \\ &= \int_0^t \frac{\lambda}{(n-1)!} \exp(-\lambda u) (\lambda u)^{n-1} \lambda \exp(-\lambda(t-u)) du \\ &= \frac{\lambda^{n+1}}{(n-1)!} \exp(-\lambda t) \int_0^t u^{n-1} du \\ &= \frac{\lambda^{n+1}}{(n-1)!} \exp(-\lambda t) \frac{t^n}{n} \\ &= \frac{\lambda}{n!} \exp(-\lambda t) (\lambda t)^n. \end{aligned}$$

Donc

$$T_n \sim \Gamma(n, \lambda).$$

■

**Propriété 2.1.3** (*Loi de  $N(t)$* )

Pour tout  $t > 0$ , la variable aléatoire  $N(t)$  représente la nombre d'événements se produisant dans l'intervalle de temps  $[0, t]$  suit la loi de Poisson de paramètre  $\lambda t$ .

**Preuve.** La démonstration repose sur l'identité

$$\{T_n \leq t\} = \{N(t) \geq n\}, \quad (t > 0, n \geq 0).$$

Alors

$$\begin{aligned} P(N(t) = n) &= P(N(t) \geq n) - P(N(t) \geq n + 1) \\ &= P(T_n \leq t) - P(T_{n+1} \leq t) \\ &= \int_0^t \lambda \exp(-\lambda x) \frac{(\lambda x)^{n-1}}{(n-1)!} dx - \int_0^t \lambda \exp(-\lambda x) \frac{(\lambda x)^n}{n!} dx \\ &= \int_0^t d \left( \exp(-\lambda x) \frac{(\lambda x)^n}{(n)!} \right) \\ &= \exp(-\lambda t) \frac{(\lambda t)^n}{(n)!}. \end{aligned}$$

Donc

$$N(t) \sim P(\lambda t).$$

■

**Propriété 2.1.4** Soient  $\{N(t), t \geq 0\}$  et  $\{M(t), t \geq 0\}$  deux processus de Poisson indépendants d'intensités respectives  $\lambda$  et  $\mu$ , alors le processus  $\{X(t), t \geq 0\}$  définie par

$$X(t) = N(t) + M(t), \quad \text{pour } t \geq 0,$$

est un processus de Poisson d'intensité  $\lambda + \mu$ .

**Remarque 2.1.1** Le résultat se généralise à la somme de  $n$  processus de Poisson.

**Propriété 2.1.5** *Un processus de Poisson est localement continue en probabilité.*

**Preuve.** Comme le processus de Poisson est à accroissements stationnaires, on a

$$\begin{aligned}
 P(N(t+h) - N(t) \geq 1) &= P(N(h) - N(0) \geq 1) \\
 &= P(N(h) \geq 1), && \text{car } N(0) = 0 \\
 &= 1 - P(N(h) < 1) \\
 &= 1 - P(N(h) = 0) \\
 &= 1 - \exp(-\lambda h) \longrightarrow 0, && \text{Lorsque } h \longrightarrow 0.
 \end{aligned}$$

■

**Propriété 2.1.6** *Soit un processus de Poisson  $\{N(t), t \geq 0\}$  d'intensité  $\lambda > 0$ ,  $\forall t$ ,  $s \geq 0$ , on a*

1.  $E[N(t)] = \text{Var}[N(t)] = \lambda t$ ,
2.  $\text{Cov}(N(s), N(t)) = \lambda \min(s, t)$ .

**Preuve.**

1. Comme

$$N(t) \sim P(\lambda t),$$

alors

$$E[N(t)] = \text{Var}[N(t)] = \lambda t.$$

2. Soit  $s < t$ , alors  $N(s)$  et  $N(t) - N(s)$  sont indépendantes (car  $N(t)$  est à accroissements indépendantes), alors

$$\text{Cov}(N(s), N(t) - N(s)) = 0,$$

dans ce cas on a

$$\begin{aligned}
 \text{Cov}(N(s), N(t)) &= \text{Cov}(N(s), N(t) - N(s) + N(s)) \\
 &= \text{Cov}(N(s), N(t) - N(s)) + \text{Cov}(N(s), N(s)) \\
 &= \text{Cov}(N(s), N(s)) \\
 &= \text{Var}[N(s)] \\
 &= \lambda s(\lambda \min(s, t)).
 \end{aligned}$$

■

### 2.1.1 Distribution conditionnelle des instants d'arrivée

#### Processus de Poisson et loi uniforme

**Propriété 2.1.7** *Conditionnement à l'évènement  $\{N(t) = n\}$ ,*

*le  $n$ -uplet  $(T_1, T_2, \dots, T_n)$  à même loi de probabilité que  $n$ -uplet ordonné correspondant à  $n$  (v.a) i.i.d de loi uniforme sur  $[0, t]$ .*

**Preuve.** Soient  $0 < x_1 < x_2 < \dots < x_n < t$  une suite strictement croissante et  $h_1, h_2, \dots, h_n$  des nombres strictement positifs suffisamment petit pour que :  $x_1 \leq x_1 + h_1 < x_2 \leq x_2 + h_2 < \dots \leq x_{n-1} + h_{n-1} < x_n \leq x_n + h_n < t$ .

Posons :  $B_n = T_1 \in [x_1, x_1 + h_1], \dots, T_n \in [x_n, x_n + h_n]$

$$\begin{aligned}
 P(B_n \mid N(t) = n) &= \frac{P(B_n, N(t) = n)}{P(N(t) = n)} \\
 &= \frac{P(\text{(un seul top dans } [x_i, x_i + h_i] (i = 1, \dots, n); 0 \text{ si non)})}{P(N(t) = n)} \\
 &= \frac{\exp(-\lambda h_1) \lambda h_1 \dots \exp(-\lambda h_n) \lambda h_n \exp(-\lambda(t - h_1 - \dots - h_n))}{\exp(-\lambda t) \frac{(\lambda t)^n}{n!}} \\
 &= n! \frac{1}{t^n} h_1 \dots h_n.
 \end{aligned}$$

On divise par  $h_1 \dots h_n$  et on fait tendre successivement  $h_1 \dots h_n$  vers 0, on trouve la densité correspondantes

$$f_{(T_1, T_2, \dots, T_n | N(t)=n)}(x_1, \dots, x_n) = n! \frac{1}{t^n} \mathbf{1}_E,$$

avec

$$E = \{(x_1, \dots, x_n) : 0 < x_1 < x_2 < \dots < x_n < t\}.$$

■

### Processus de Poisson et loi binomiale

**Propriété 2.1.8** Pour  $s < t$ , La loi conditionnelle de  $N_s$  sachant  $\{N(t) = n\}$  est la loi binomiale  $B(n, \frac{s}{t})$ .

**Preuve.**  $\forall s, t (s < t), \forall k \leq n$

$$\begin{aligned} P(N(s) = k | N(t) = n) &= \frac{P(N(s) = k, N(t) = n)}{P(N(t) = n)} \\ &= \frac{P(N(s) = k, N(t) - N(s) = n - k)}{P(N(t) = n)} \\ &= \frac{P(N(s) = k)P(N(t) - N(s) = n - k)}{P(N(t) = n)} \\ &= \frac{\exp(-\lambda s) \frac{(\lambda s)^k}{k!} \exp(-\lambda(t-s)) \frac{(\lambda(t-s))^{n-k}}{(n-k)!}}{\exp(-\lambda t) \frac{(\lambda t)^n}{n!}} \\ &= C_n^k \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}. \end{aligned}$$

■

**Exemple 2.1.2** *Le nombre de clients arrivant dans une épicerie peut être modélisé par un processus de Poisson d'intensité  $\lambda = 4$  clients par heure. Quand l'épicerie ouvre à 9 : 00 A.M. Quelle est la probabilité qu'un seul client soit arrivé à 9 : 30 A.M et que cinq au total soient arrivés à 11 : 30 A.M ?*

*On calcule les longeuere des intervalles de temps  $t$  en heures à partir de 9 : 00 A.M, alors*

- *L'intevalle de temps entre 9 : 00 A.M et 9 : 30 A.M a une longeuere  $\tau_1 = \frac{1}{2}$  heures.*
- *L'intevalle de temps entre 9 : 00 A.M et 11 : 30 A.M a une longeuere  $\tau_1 = \frac{5}{2}$  heures.*

*Et on nous demande de déterminet  $P(X(\frac{1}{2}) = 1, X(\frac{5}{2}) = 5)$ .*

*On utilise l'indépendance de  $X(\frac{5}{2}) - X(\frac{1}{2})$  et  $X(\frac{1}{2})$  pour reformuler la question :*

$$\begin{aligned}
 P\left(X\left(\frac{1}{2}\right) = 1, X\left(\frac{5}{2}\right) = 5\right) &= P\left(X\left(\frac{1}{2}\right) = 1, X\left(\frac{5}{2}\right) - X\left(\frac{1}{2}\right) = 4\right) \\
 &= P\left(X\left(\frac{1}{2}\right) = 1\right) P\left(X\left(\frac{5}{2}\right) - X\left(\frac{1}{2}\right) = 4\right) \\
 &= \left[ \frac{\exp\left(-4\left(\frac{1}{2}\right) 4\left(\frac{1}{2}\right)\right)}{1!} \right] \left[ \frac{\exp(-4(2)) [4(2)]^4}{4!} \right] \\
 &= 0.0154965.
 \end{aligned}$$

## 2.2 Estimation par la méthode du maximum de vraisemblance

La méthode du maximum de vraisemblance est une méthode statistique qui permet d'estimer les paramètres d'un modèle. La technique consiste à construire une fonction appelée fonction de vraisemblance et à maximiser son logarithme par rapport aux paramètres inconnus. Les estimateurs obtenus par cette méthode ont de bonnes propriétés statistiques.

Soit  $X$  une variable aléatoire réelle de loi paramétrique (discrète ou continue), dont on veut estimer le paramètre inconnu  $\theta$ . Alors on définit une fonction  $f$  telle que

$$f(x; \theta) = \begin{cases} f_{\theta}(x), & \text{si } X \text{ continue de densité } f \\ P_{\theta}(X = x), & \text{si } X \text{ discrète de probabilité } P. \end{cases}$$

**Définition 2.2.1** On appelle fonction de vraisemblance de  $\theta$  pour une réalisation  $(x_1, \dots, x_n)$  d'un échantillon, la fonction de  $\theta$  définie par

$$L(x_1, \dots, x_n; \theta) = f(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta).$$

**Définition 2.2.2** On appelle estimateur du maximum de vraisemblance de  $\theta$  (notée par EMV) pour une réalisation  $(\mathbf{x}_1, \dots, \mathbf{x}_n)$  un réel  $\hat{\theta}$  qui maximise la fonction de vraisemblance  $L(x_1, \dots, x_n; \theta)$  en  $\theta$ , i.e. pour tout  $\theta$

$$L(x_1, \dots, x_n; \theta) \leq L(x_1, \dots, x_n; \hat{\theta}),$$

une expression alternative est

$$\hat{\theta} \in \arg \max_{\theta} L(x_1, \dots, x_n; \theta).$$



Pour obtenir l'expression de l'estimateur du maximum de vraisemblance  $\hat{\theta}$  de  $\theta$  pour  $(x_1, \dots, x_n)$ , une idée est d'exprimer  $L(x_1, \dots, x_n; \theta)$  en fonction de produits de termes exponentiels /puissances, puis de considérer la fonction de log-vraisemblance pour  $(x_1, \dots, x_n)$  fonction de  $\theta$  définie par

$$l(x_1, \dots, x_n; \theta) = \log(L(x_1, \dots, x_n; \theta)),$$

elle n'a de sens que si  $\theta$  vérifie  $L(x_1, \dots, x_n; \theta) > 0$ , si cette dernière (fonction log-vraisemblance) est dérivable en  $\theta$  une condition nécessaire que doit vérifier  $\hat{\theta}$  est d'être solution de l'équation de vraisemblance

$$\frac{d}{d\theta} (l(x_1, \dots, x_n; \theta)) = 0,$$

i.e. la solution  $\theta$  de cette équation est l'EMV  $\hat{\theta}$ .

**Exemple 2.2.1** *On souhaite estimer le paramètre  $\lambda$  d'une loi de Poisson à partir d'un  $n$ -échantillon .on a*

$$f(x; \lambda) = P_\lambda(X = x) = \frac{\lambda^x}{x!} \exp(-\lambda),$$

la fonction de vraisemblance s'écrit

$$\begin{aligned} L(x_1, \dots, x_n; \lambda) &= \prod_{i=1}^n f(x_i; \lambda) \\ &= \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} \exp(-\lambda) \\ &= \exp(-\lambda n) \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!}. \end{aligned}$$

Comme la fonction vraisemblance étant positive, alors on peut utiliser le logarithme

$$\begin{aligned}
 l(x_1, \dots, x_n; \lambda) &= \log(L(x_1, \dots, x_n; \lambda)) \\
 &= \log(\exp(-\lambda n)) + \log\left(\prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!}\right) \\
 &= -\lambda n + \sum_{i=1}^n \log\left(\frac{\lambda^{x_i}}{x_i!}\right) \\
 &= -\lambda n + \log(\lambda) \sum_{i=1}^n x_i - \sum_{i=1}^n \log(x_i!).
 \end{aligned}$$

La dérivée première

$$\frac{d}{d\lambda} (l(x_1, \dots, x_n; \lambda)) = -n + \frac{1}{\lambda} \sum_{i=1}^n x_i.$$

S'annule pour

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Alors l'EMV de paramètre  $\lambda$  d'une loi de Poisson à partir d'un  $n$ - échantillon est

$$\hat{\lambda} = \bar{X}.$$

Donc comment estimer l'intensité de processus de Poisson par cette méthode ?

## 2.3 Estimation de l'intensité d'un processus de Poisson homogène

Soit un processus de Poisson  $\{N(t), t \geq 0\}$  d'intensité  $\lambda$  inconnue, on se propose d'estimer cette densité à partir d'observations. Donnons deux méthodes

**Première méthode :** Supposons qu'on observe le processus jusqu'à un instant  $t > 0$ , on dispose alors des données suivantes

a/  $n$ , le nombre d'événements dans  $[0, t]$ ;

b/ Les instants  $x_1, \dots, x_n$  ( $0 < x_1 < \dots < x_n < t$ ), où se produisent les  $n$  événements consécutifs dans  $[0, t]$ .

$$\begin{aligned} L(n, x_1, \dots, x_n; \lambda) &= P(N(t) = n) g(x_1, \dots, x_n) \\ &= \exp(-\lambda t) \frac{(\lambda t)^n}{n!} \left(\frac{n!}{t^n}\right) (0 < x_1 < x_2 < \dots < x_n < t) \\ &= \lambda^n \exp(-\lambda t), \end{aligned}$$

où  $g(x_1, \dots, x_n)$  désigne la densité de la loi de  $(T_1, \dots, T_n \mid N_t = n)$  donnée par

$$g(x_1, \dots, x_n) = \frac{(\lambda t)^n}{n!} \mathbf{1}_{(0 < x_1 < x_2 < \dots < x_n < t)},$$

et les temps d'arrivées  $T_1, \dots, T_n$  déterminent la trajectoire du processus sur  $[0, t]$ .

Alors

$$\begin{aligned} l(n, x_1, \dots, x_n; \lambda) &= \log(L(n, x_1, \dots, x_n; \lambda)) \\ &= n \log(\lambda) - \lambda t. \end{aligned}$$

La dérivée première est

$$\frac{d}{d\lambda} (l(n, x_1, \dots, x_n; \lambda)) = \frac{n}{\lambda} - t.$$

S'annule pour

$$\hat{\lambda} = \frac{n}{t}.$$

Ainsi l'estimateur par la méthode du maximum de vraisemblance est

$$\hat{\lambda} = \frac{n}{t}.$$

L'estimateur correspondant est obtenue en remplaçant  $n$  par  $N(t)$ , et on obtient

$$\hat{\lambda} = \frac{N(t)}{t}.$$

**Proposition 2.3.1** *L'estimateur  $\hat{\lambda}$  est non biaisé, exhaustif et complet.*

**Preuve.** En effet

$$\begin{aligned} E[\hat{\lambda}] &= E\left[\frac{N(t)}{t}\right] \\ &= \frac{1}{t}E[N(t)] \\ &= \frac{1}{t}(\lambda t) = \lambda. \end{aligned}$$

Comme la fonction de vraisemblance  $L(n, x_1, \dots, x_n; \lambda)$  est de la forme  $\lambda^n \exp(-\lambda t)$ , l'estimateur est bien exhaustif. Et on sait que  $N(t)$  suit une loi de Poisson de paramètre  $\lambda t$  et que la loi de Poisson est complète, donc l'estimateur  $\hat{\lambda}$  est complet.

■

**Remarque 2.3.1** *Puisque  $\hat{\lambda}$  est non biaisé, exhaustif et complet, alors  $\hat{\lambda}$  est l'unique estimateur non biaisé de variance minimum de  $\lambda$ .*

### 2.3.1 Comportement asymptotique

**Théorème 2.3.1** *Au niveau du comportement asymptotique de  $N(t)$ , on a les deux résultats suivantes*

1.  $\lim_{t \rightarrow +\infty} \frac{N(t)}{t} = \lambda$  p.s.  
(i.e l'estimateur  $\hat{\lambda}$  est un estimateur consistant car  $\hat{\lambda} = \frac{N(t)}{t} \xrightarrow[t \rightarrow +\infty]{p.s.} \lambda$ ).
2. On a de plus la convergence en loi

$$\sqrt{\frac{t}{\lambda}} \left( \frac{N(t)}{t} - \lambda \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1), \quad \text{Lorsque } t \longrightarrow +\infty.$$

**Preuve.**

1. Notons  $X_1, X_2, \dots, X_{[t]}$  le nombre d'occurrences d'événements dans les intervalles  $[0, 1], [1, 2], \dots, [t-1, t]$  ( $t > 1$ ) où  $[t]$  est la partie entière de  $t$ . Les  $X_k$  sont des variables aléatoires, indépendantes, identiquement distribuées, d'espérance mathématique  $\lambda$ . D'après la loi forte des grands nombres, on a

$$\frac{N([t])}{[t]} = \frac{X_1 + X_2 + \dots + X_{[t]}}{[t]} \xrightarrow{p.s} E[X_1] = \lambda,$$

et de même

$$\frac{N(t)}{t} \xrightarrow[t \rightarrow +\infty]{p.s} \lambda.$$

2. On utilise le théorème de la limite central et la décomposition :

$$\sqrt{\frac{t}{\lambda}} \left( \frac{N(t)}{t} - \lambda \right) = \sqrt{\frac{t}{\lambda}} \left( \frac{N([t])}{[t]} - \lambda \right) + A_t.$$

On pose

$$N([t]) = \sum_{i=1}^{[t]} (N(i) - N(i-1)),$$

est la somme de  $n$  (v.a) indépendantes, de même loi de Poisson de paramètre  $\lambda$ .

- Pour le premier terme on a

$$\sqrt{\frac{t}{\lambda}} \left( \frac{N([t])}{[t]} - \lambda \right) = \sqrt{\frac{t}{[t]}} \sqrt{\frac{[t]}{\lambda}} \left( \frac{\sum_{i=1}^{[t]} (N(i) - N(i-1))}{[t]} - \lambda \right).$$

On obtient alors la convergence en loi de ce premier terme vers la loi  $\mathcal{N}(0, 1)$  on utilise le théorème central limite et le lemme de Slutsky  $\left( \lim_{t \rightarrow +\infty} \frac{t}{[t]} = 1 \right)$ .

- Pour le deuxième terme, on a

$$\begin{aligned} |A_t| &= \sqrt{\frac{t}{\lambda}} \left| \frac{N([t])}{[t]} - \frac{N(t)}{t} \right| \\ &\leq \sqrt{\frac{t}{\lambda}} \left( \frac{N(t) - N([t])}{[t]} + \frac{N(t)}{[t]t} \right). \end{aligned}$$

On pourra montrer que le premier de deux termes majorants tend vers 0 en probabilité on utilise l'inégalité de Markov, comme le deuxième tend vers 0 *p.s* on utilise le premier point du théorème, on conclut que  $\lim_{t \rightarrow +\infty} A_t = 0$  en probabilité. D'après ce qui précède et on utilise le lemme de Slutsky, on obtient la convergence. ■

**Remarque 2.3.2** On déduit que l'intervalle de confiance (asymptotique) pour  $\lambda$  au niveau  $1 - \alpha$  est

$$\mathbf{I}_n = \left[ -\sqrt{\frac{N(t)}{t^2}} q_{1-\frac{\alpha}{2}} + \frac{N(t)}{t}, +\sqrt{\frac{N(t)}{t^2}} q_{1-\frac{\alpha}{2}} + \frac{N(t)}{t} \right],$$

où  $q_{1-\frac{\alpha}{2}}$  est le quantile d'ordre  $1 - \frac{\alpha}{2}$  de la loi  $\mathcal{N}(0, 1)$ .

**Deuxième méthode :** supposons qu'on observe le processus jusqu'à l'apparition du  $n^{\text{ième}}$  évènement. On dispose alors la donnée des instants  $x_1, \dots, x_n$  des apparitions des évènements. La fonction de vraisemblance est donnée par

$$\begin{aligned} L(n, x_1, \dots, x_n; \lambda) &= \lambda \exp(-\lambda x_1) \dots \lambda \exp(-\lambda x_n) \\ &= \lambda^n \exp(-\lambda(x_1 + \dots + x_n)). \end{aligned}$$

Alors

$$\begin{aligned} l(n, x_1, \dots, x_n; \lambda) &= \log(L(n, x_1, \dots, x_n; \lambda)) \\ &= n \log(\lambda) - \lambda(x_1 + \dots + x_n). \end{aligned}$$

Par suite

$$\frac{d}{d\lambda} (l(n, x_1, \dots, x_n; \lambda)) = \frac{n}{\lambda} - (x_1 + \dots + x_n).$$

S'annule pour

$$\hat{\lambda} = \frac{n}{(x_1 + \dots + x_n)}.$$

On pose

$$T_n = X_1 + \dots + X_n,$$

instant d'arrivée du  $n^{\text{ième}}$  évènement, on voit que l'estimateur correspondant est

$$\hat{\lambda} = \frac{n}{T_n}.$$

Au niveau du comportement asymptotique de  $N(t)$ , on a les deux résultats suivantes

**Théorème 2.3.2** *On a*

1.  $\hat{\lambda} = \frac{n}{T_n} \xrightarrow{p.s} \lambda$ .
2.  $\sqrt{n} \left( \lambda \frac{T_n}{n} - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ .

# Conclusion

L'étude des processus ponctuels est l'un des sujets centraux des processus stochastiques et a été largement utilisée pour modéliser des événements discrets en temps continu.

En particulier, le processus de Poisson, un processus ponctuel commun, a le plus d'applications. Les exemples classiques incluent les arrivées de clients dans un parc sur une période de temps, les buts marqués lors d'un match de football d'association et les clics sur un lien Web particulier au cours d'une période donnée. Récemment, les processus de Poisson ont été utilisés pour caractériser l'activité de pointe dans divers systèmes neuronaux.

Dans ce mémoire, on a défini le processus de Poisson homogène et ces propriétés fondamentales, puis on a estimé l'intensité  $\lambda$  de ce processus par la méthode de maximum de vraisemblance et on a étudié le comportement asymptotique de cet estimateur.

Lorsque l'intensité  $\lambda$  n'est pas une constante, c'est-à-dire devient une fonction du temps (cas d'un processus de Poisson non homogène), la question sera comment on peut estimer cette intensité ?



# Bibliographie

- [1] Breton, J. CH. (2020). Processus stochastiques. Université de Rennes 1.
- [2] Caumel, Y. (2011). Probabilités et processus stochastiques. Springer.
- [3] Duheille-Bienvenue, F.(2006). Processus stochastiques. Université de Claude Bernars Lyon 1.
- [4] Dusart, P.(2018).cours de statistique inférentielles.
- [5] Foata, D. & Fuchs, A. (2004). Processus stochastique : Processus de Poisson, chaîne de Markov et martingales. Edition Dunod.
- [6] Goffard, P. O. (2018). Processus de Poisson. Université de Lyon 1
- [7] Jacod, J. (2003). Chaînes de Markov, processus de Poisson et applications. DEA de probabilités et applications, 2004.
- [8] Khaled, K.(2010). Probabilités. OPU.
- [9] Lefebvre, M. (2007). Applied stochastic processes. Springer Science & Business Media.
- [10] Lepetit, G. Processus de Poisson. ENS Rennes-Université Rennes 1.
- [11] Ross, Sheldon M., et al.(1996). Stochastic processes. New York : Wiley.
- [12] Ross, S. M. (2014). Introduction to probability models. Academic press.

# Annexe A : Logiciel R

Le langage R est un logiciel dans lequel de nombreuses techniques statistiques modernes et classiques ont été implémentées. Il comporte des moyens qui rendent possible la manipulation des données, les représentations graphiques et les calculs.

Dans ce mémoire on va donner les représentations graphiques de quelque lois de probabilités et on va simuler le processus de Poisson homogène.

## Loi de Poisson

```
par(mfrow=c(2,2))
```

```
plot(0 :20,dpois(0 :20,0.75),xlab="k",ylab="P(x=k)",type="h",main="Poisson pour  
lambda=0.75")
```

```
plot(0 :20,dpois(0 :20,3),xlab="k",ylab="P(x=k)",type="h",main="Poisson pour  
lambda=3")
```

```
plot(0 :20,dpois(0 :20,6),xlab="k",ylab="P(x=k)",type="h",main="Poisson pour  
lambda=6")
```

```
plot(0 :20,dpois(0 :20,10),xlab="k",ylab="P(x=k)",type="h",main="Poisson pour  
lambda=10")
```

**Loi exponentielle**

```
curve(dexp(x, 0.5), xlim = c(0, 3), ylim = c(0, 1), col = "red", lwd = 2,  
ylab = "f(x)", main = "Exemples de densités associées à la loi exponentielle")  
curve(dexp(x, 1), col = "green", lwd = 2, add = TRUE)  
curve(dexp(x, 1.5), col = "blue", lwd = 2, add = TRUE)  
curve(dexp(x, 2), col = "black", lwd = 2, add = TRUE)  
legend("topright", legend = c("E(0.5)", "E(1)", "E(1.5)", "E(2)"),  
lty = 1, lwd = 1, col = c("red", "green", "blue", "black"))
```

**Processus de Poisson homogène**

```
par(mfrow=c(1,2))  
plot(stepfun(cumsum(rexp(10,3)),  
cumsum(c(0,rep(1,10))))),do.points=T,xlab="Temps",  
ylab="",main="Trajectoire d'un processus de poisson d'intensité lambda=3",  
lw=2,col="blue")  
plot(stepfun(cumsum(rexp(10,7)),cumsum(c(0,rep(1,10))))),do.points=T,xlab="Temps",  
ylab="",main="Trajectoire d'un processus de poisson d'intensité lambda=7",  
lw=2,col="blue")
```

# Annexe B : Abréviations et Notations

Les différentes abréviations et notations utilisées tout au long de ce mémoire sont expliquées ci-dessous :

$E(.)$	: Espérance mathématique.
$V(.)$	: Variance mathématique.
$Cov(X, Y)$	: Covariance mathématique du couple $(X, Y)$ .
$P$	: Probabilité
$F_X$	: Fonction de répartition de $X$ .
$f_X$	: Densité de probabilité de $X$ .
$(\Omega, \mathcal{F}, P)$	: Espace de probabilité.
$(\mathbb{E}, \mathfrak{B}(\mathbb{E}))$	: Espace mesurable.
$\mathcal{P}(\lambda)$	: Loi de Poisson de paramètre $\lambda$ .
$\mathcal{E}(\lambda)$	: Loi exponentielle de paramètre $\lambda$ .
$\mathcal{N}(0, 1)$	: Loi normal centrée réduite.
$\Gamma(n, \lambda)$	: Loi Gamma de paramètre $n$ et $\lambda$ .
$\{N(t), t \geq 0\}$	: Processus de comptage (ou processus de Poisson).
$iid$	: indépendantes et identiquement distribuées.
$\stackrel{\mathcal{L}}{=}$	: L'égalité en loi.
$\stackrel{\mathcal{L}}{\rightarrow}$	: L'approximation en loi.

## Abstract

**F**or the homogeneous Poisson process there are already common methods for estimating the intensity, which is a constant in the homogeneous case, for example, maximum likelihood method, moment, ect.

We deal here with the estimation of the intensity of a homogeneous Poisson process by the maximum likelihood method.

**Keywords :** Homogeneous Poisson process , Intensity, Maximum likelihood method .

## Résumé

**P**our le processus de Poisson homogène, il y a déjà des méthodes communes pour estimer l'intensité, qui est une constante dans le cas homogène. Par exemple, méthode du maximum de vraisemblance, méthode de moment, ect.

Nous traitons ici l'estimation de l'intensité d'un processus de Poisson homogène par la méthode du maximum de vraisemblance.

**Mots-clés :** Processus de Poisson homogène, Intensité, méthode du maximum de vraisemblance.

## ملخص

من اجل عملية بواسون، توجد عدة طرق لتقدير الكثافة، و هي ثابتة في الحالة المتجانسة، على

سبيل المثال، طريقة الاحتمالية العظمى، طريقة اللحظة، الخ.

في هذا العمل قمنا بتقدير كثافة عملية بواسون المتجانسة بطريقة الاحتمالية العظمى.

**الكلمات المفتاحية:** عملية بواسون المتجانسة، الكثافة، طريقة الاحتمالية العظمى.