

# PROCESSUS PONCTUELS MARKOVIENS

François JAULIN

Juin 2008

## Résumé

Dans ce qui suit, on rappelle dans un premier temps quelques notions générales sur les processus ponctuels. On introduit ensuite les processus ponctuels de Poisson, à partir desquels on définit une variété de processus, les processus à densité. Enfin, on présente les processus ponctuel de Markov, très utilisés en pratique. Ce cheminement est largement inspiré du livre de M.N.M. van Lieshout, intitulé Markov Point Process and their Applications

## Table des matières

<b>1</b>	<b>Processus ponctuels : généralités</b>	<b>3</b>
1.1	Introduction et motivations . . . . .	3
1.2	Définition d'un processus ponctuel . . . . .	6
1.3	Fidis . . . . .	8
1.4	Processus ponctuels marqués . . . . .	8
1.5	Processus ponctuels simples . . . . .	9
1.6	Processus ponctuels finis . . . . .	10
1.6.1	Caractérisation d'un processus ponctuel fini . . . . .	10
1.6.2	Exemple fondamental : le processus binomial . . . . .	11
<b>2</b>	<b>Processus de Poisson</b>	<b>12</b>
2.1	Processus de Poisson . . . . .	12
2.1.1	Processus de Poisson homogène . . . . .	12
2.1.2	Processus ponctuel de Poisson inhomogène . . . . .	13
2.1.3	Processus ponctuel de Poisson . . . . .	14
2.1.4	Commentaires . . . . .	14
<b>3</b>	<b>Processus fini définit à l'aide d'une densité</b>	<b>15</b>
3.1	Exemples introductifs : le processus agrégé et le modèle booléen . . . . .	15
3.1.1	Le processus de Poisson agrégé : modèle de Neymann-Scott . . . . .	16
3.1.2	Le modèle booléen . . . . .	16
3.2	Définition et exemples . . . . .	17
3.2.1	Définition . . . . .	17
3.2.2	Exemples . . . . .	17
<b>4</b>	<b>Processus ponctuel de Markov</b>	<b>18</b>
4.1	Exemple introductif et rappels sur les chaines de Markov . . . . .	18
4.1.1	Le modèle des sphères dures . . . . .	18
4.1.2	Rappels sur les chaines de Markov . . . . .	19
4.2	Notions de voisinage . . . . .	19
4.3	Densité conditionnelle de Papangelou . . . . .	20
4.3.1	Interprétation de la densité conditionnelle de Papangelou . . . . .	20
4.3.2	Une expression de la densité conditionnelle de Papangelou . . . . .	20
4.4	Processus ponctuel de Markov . . . . .	20
4.5	Théorème et propriétés fondamentales . . . . .	21
4.5.1	La markovianité spatiale . . . . .	21
4.5.2	Le théorème d'Hammersley - Clifford . . . . .	21
4.6	Exemples . . . . .	22
4.6.1	Processus ponctuel de poisson d'intensité constante . . . . .	22

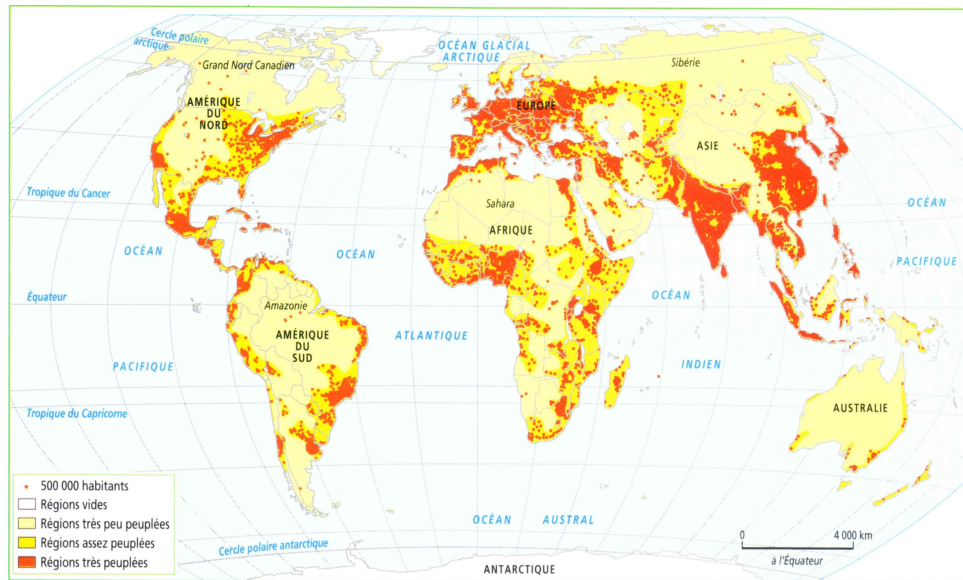
4.6.2	Le modèle des shères dures . . . . .	23
4.6.3	Cas général : le modèle des paires interactives . . . . .	23
<b>5</b>	<b>Processus Ponctuels Markoviens Marqués</b>	<b>23</b>
5.1	Définitions . . . . .	23
5.2	Exemples . . . . .	24
<b>6</b>	<b>Conclusion</b>	<b>24</b>

# 1 Processus ponctuels : généralités

## 1.1 Introduction et motivations

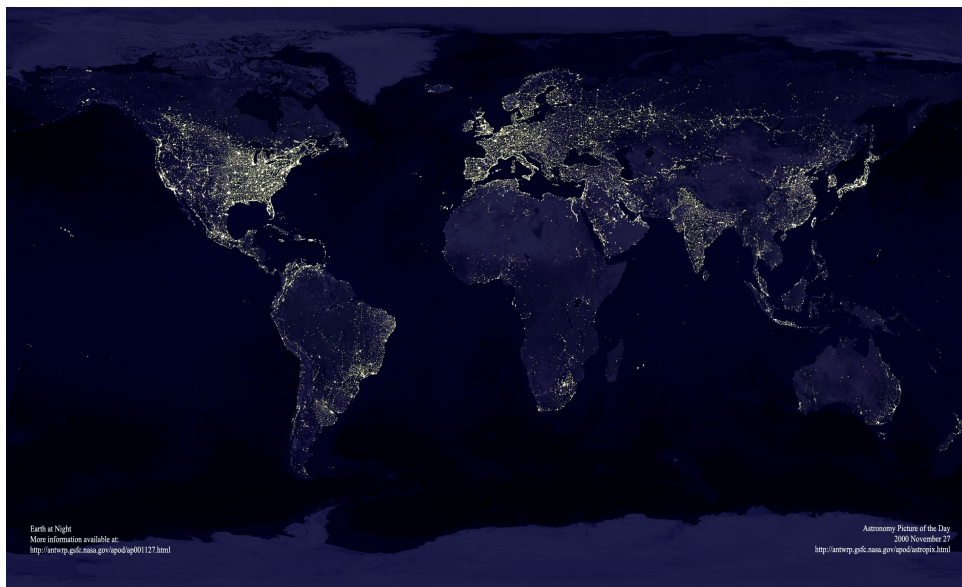
De nombreuses expériences, issues de différents domaines, fournissent des données ponctuelles.

Par exemple, en démographie, lorsqu'on s'intéresse à la répartition de la population humaine sur le globe, on peut à l'aide d'outils de recensement et d'outils statistiques fournir des résultats sous forme ponctuelle. Sur l'image suivante, chaque point représente 500 000 habitants.



Repartition de la population humaine sur la Terre en 2005

L' image suivante, devenue célèbre pour sa beauté et son originalité - photographier par satellite la Terre pendant la nuit - présente aussi la caractéristique de donner une répartition de points (blancs) associés aux sources de lumière créées par l'homme.



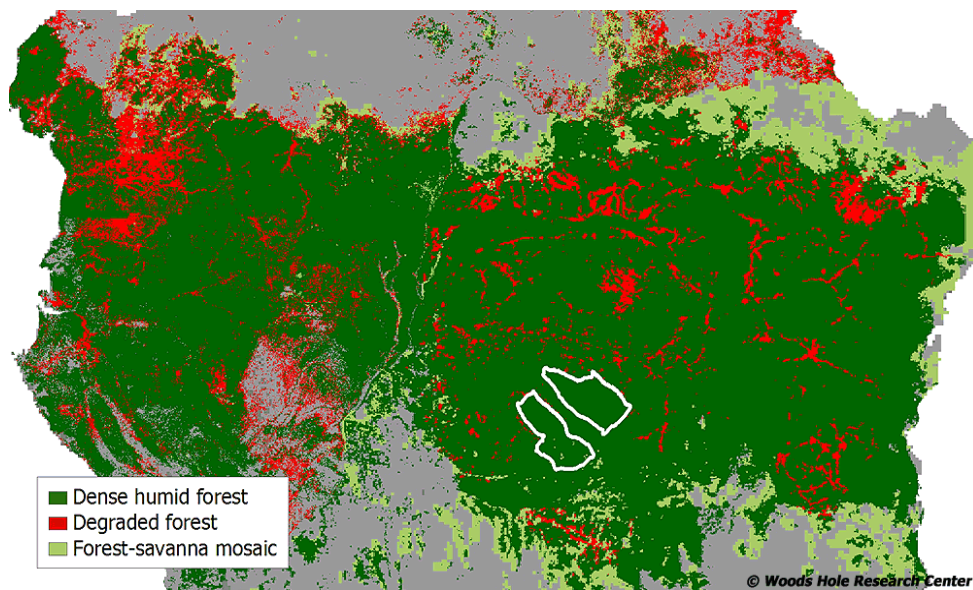
La Terre vue de l'espace la nuit

Ceci donne une idée assez vague de la répartition de la population humaine en termes de villes - les points lumineux correspondent aux rayonnements des villes dans lesquelles les hommes sont concentrés.

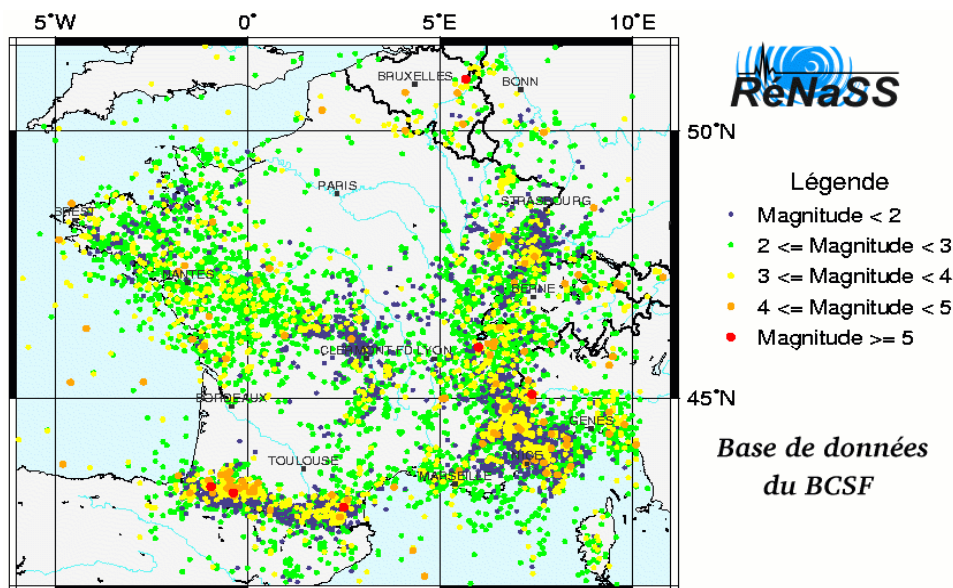
On peut par exemple noter que les États Unis sont très lumineux pour 300 millions d'habitants comparé à la Chine et ses 1 307 millions d'habitants, et que l'Afrique ou des pays comme la Corée du Nord sont "éteints".

Sur le plan ponctuel, on remarque que les points sont rarements isolés et semblent s'attirer. Ils forment des taches, poreuses par endroits (Europe, Ouest des États Unis, ...), à cavité (Sud de l'Afrique, Australie, Asie Centrale, ...), éclatées (Afrique Centrale, Nord de l'Amérique du Sud, Alaska, ...).

La démographie n'est pas le seul domaine dans lequel les points fourmillent. La géographie ainsi que les matières qui s'y rapportent (géologie, agriculture, sismologie ...) fournissent une multitude de données ponctuelles.



Carte de la déforestation (points rouges) au parc national de la Salonga - whcr.com



Épicentres et zone de sismicité en France - base de donnée BCSF

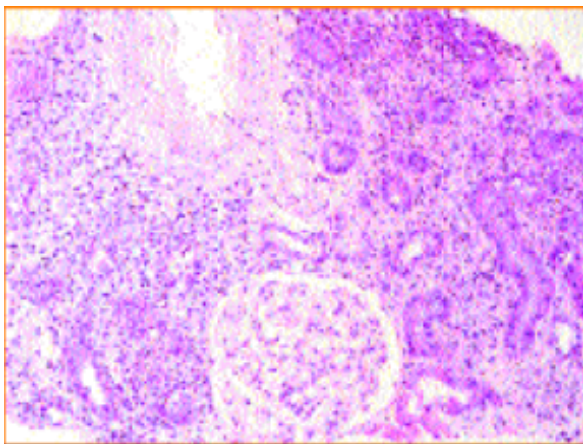
Les points sont aussi très étudiés par les astronomes et astrophysiciens. Nous verrons plus loin que certains modèles furent créé pour représenter les amas de galaxies.



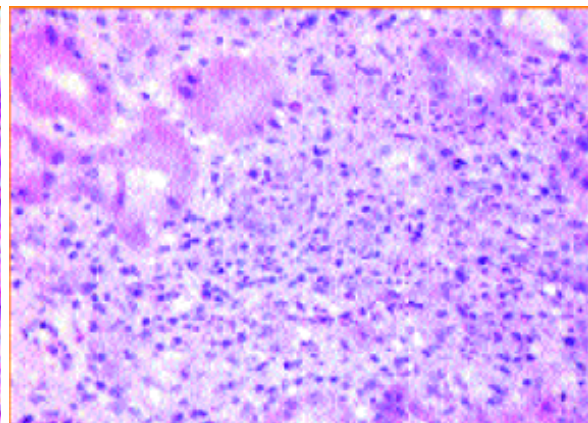


Galaxie Pléiades

En biologie, épidémiologie, ou tout autre domaine qui s'attache à étudier un aspect des cellules vivantes, les résultats sont souvent aussi sous formes de données ponctuels.



**Fig. 1. Néphrite tubulo-interstitielle diffuse, sévère, composée principalement de polymorphonucléaires.**  
Glomérule sans lésion notable. Fibroélastose artérielle modérée. Coloration hématoxyline-éosine. Grossissement 10 X.



**Fig. 2. Infiltrat inflammatoire interstitiel sévère, composé de polymorphonucléaires, de quelques histiocytes, plasmocytes et de rares éosinophiles.**  
Infiltration des cellules épithéliales tubulaires par les polymorphonucléaires avec nécrose focale. Coloration hématoxyline-éosine. Grossissement 40 X.

Tuberculose - Revue Médicale Suisse N° – 574

Sans les analyser chacune précisément - ce n'est pas notre objectif ici - on peut déjà remarquer que les images qui précèdent fournissent des répartitions de points qui ont souvent tendance à découper l'espace en plusieurs régions distinctes. Parfois, la répartition des points révèle certaines propriétés intéressantes sur le plan expérimental : taches de points attractives (cas des villes) ou répulsives (cas d'espèces animales formant un modèles prédateur-proie, ...), interaction à courte distance (cellules vivantes, ...), consistance des taches, poreuse, à cavité, ...

Ces images motivent donc à formuler un modèle mathématique qui pourrait permettre de comprendre ou d'interpréter ces répartitions de points. Rappelons ici que ces images présentent du point de vue probabiliste/statistique une réalisation du hasard et on appelle processus ponctuels les applications qui fournissent en fonction du hasard ces répartitions de points.

En vue d'étudier un certain type de processus, les processus ponctuels markoviens qui permettent de prendre en compte l'environnement et les dépendances éventuelles entre les différents points des expériences, donnons à

présent les définitions et les propositions qui vont nous permettre de modéliser et décrire plus précisément ces répartitions de points.

Certains points ne seront pas approfondis, en particulier certaines démonstrations, qui n'apportent pas d'éléments particulièrement intéressants. Le lecteur pourra s'il le souhaite les retrouver soit directement dans le livre *Markov Point Processes and their Applications*, soit indirectement dans les références proposé par van LIESHOUT.

## 1.2 Définition d'un processus ponctuel

Soit  $(\Omega, A, \mathbb{P})$  un espace probabilisé. Cet espace caractérisera les aspects aléatoires des expériences.

Soit  $\chi$  un ensemble non vide. Cet ensemble servira de contenant. En vue d'utiliser les résultats de la théorie de la mesure, on va munir  $\chi$  d'une distance qui en fera un espace métrique complet séparable. Dans la plupart des cas envisagés ici,  $\chi$  sera égale à  $(\mathbb{R}^p, \|\cdot\|_2)$  ou un compact de cet espace et  $\mu$  sera la mesure de Lebesgue sur cet espace muni de sa tribu borélienne (tribu engendrée par les ouverts  $\chi$ ).

**Définition 1 (Configuration).** On appelle configuration tout ensemble de points  $\mathbf{x} = (x_1, \dots, x_n, \dots)$  où  $x_n \in (\chi, d)$  sont des points issus d'une expérience aléatoire. On dit qu'une configuration est localement finie si elle possède au plus un nombre fini de points dans n'importe quel borélien borné  $A$  de  $(\chi, d)$ . On note  $N^{lf}$  la famille des configurations localement finies.

Une configuration localement finie  $\mathbf{x}$  ne possède donc pas de points d'accumulation :

$$\forall x \in \chi, \forall r > 0, \mathbf{x} \cap B(x, r) \text{ est de cardinal fini.}$$

**Définition 2 (Processus ponctuels).** On appelle processus ponctuel sur  $(\chi, d)$  une application qui à une réalisation possible du hasard  $\omega \in (\Omega, A, P)$  associe une configuration localement finie  $\mathbf{x} \in N^{lf}$  dans  $\chi$  dont le nombre de points  $N(A)$  dans un borélien borné  $A$  vérifie :  $\omega \mapsto N(A)(\omega) = N(A)$  est une variable aléatoire discrète finie pour tout borélien  $A$  borné de  $\chi$ . (voir Figure 1)

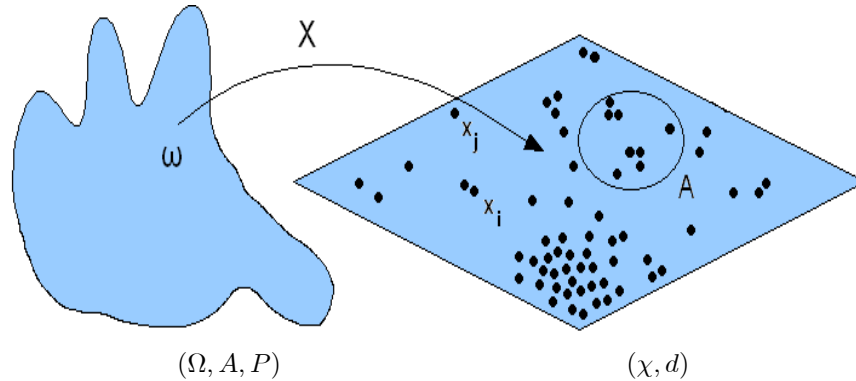


Figure 1

**Exemple 1.** Le premier exemple qui peut nous venir à l'esprit est celui d'un processus ponctuel qui envoie de manière uniforme un point dans un segment :  $\chi = [0, a]$  et  $X = \{U\}$  où  $U : \begin{matrix} \Omega \rightarrow [0, a] \\ \omega \mapsto u \end{matrix}$  suit une loi uniforme. Dans ce cas :  $N(A) = \mathbf{1}(u \in A)$ .

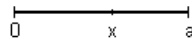


Figure 2

**Exemple 2.** Regardons le processus  $X = \{X_1, X_1 + X_2, X_1 + X_2 + X_3, \dots\}$  où  $\{X_i : \Omega \rightarrow \mathbb{R}^+\}_{i \in \mathbb{N}^*}$  désigne une suite de loi exponentielle de paramètre  $\lambda$  à termes indépendants. Un tel processus est de Poisson, d'intensité  $\lambda$ .

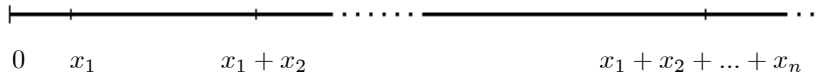


Figure 3

En effet, pour tout borélien  $A$  de  $\mathbb{R}^+$  borné,  $N(A) = \sum_{k=1}^{\infty} \mathbf{1}(X_1 + X_2 + \dots + X_k \in A)$  est une variable aléatoire en tant que limite de variable aléatoire. Montrons que  $N(A)$  est p.s. fini.

Prenons un intervalle  $[0, t]$  qui contient  $A$  - existe puisque  $A$  est borné - et regardons  $N([0, t]) \geq N(A)$  :

$$\begin{aligned}
 \mathbb{P}(N([0, t]) = k) &= \mathbb{P}(X_1 + X_2 + \dots + X_k \leq t; X_1 + X_2 + \dots + X_k + X_{k+1} > t) \\
 *^1 &= \int_{\mathbb{R}} \mathbf{1}(x \leq t) \mu_{X_1 + X_2 + \dots + X_k}(dx) \int_{\mathbb{R}} \mathbf{1}(y > t - x) \mu_{X_{k+1}}(dy) \\
 &= \int_{\mathbb{R}} \mathbf{1}(x \leq t) \mu_{X_1 + X_2 + \dots + X_k}(dx) \int_{t-x}^{+\infty} \lambda e^{-\lambda y} dy \\
 *^2 &= \int_0^t \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x} e^{-\lambda(t-x)} dx \\
 &= \frac{(\lambda t)^k}{k!} e^{-\lambda t}
 \end{aligned}$$

On passe de la ligne 1 à la ligne 2, repérée par  $*^1$ , en remarquant que :

$$\mathbf{P}(X_1 + X_2 + \dots + X_k \leq t; X_1 + X_2 + \dots + X_k + X_{k+1} > t) = \mathbf{P}(X \leq t; X + Y > t)$$

avec  $X = X_1 + X_2 + \dots + X_k$  et  $Y = X_{k+1}$  qui sont des variables aléatoires indépendantes.

Le passage de la ligne 3 à la ligne 4, repéré par  $*^2$ , est obtenu en utilisant la stabilité par l'addition des lois gamma que l'on rappelle : la somme de deux lois gamma indépendantes  $\gamma(k_1, \lambda)$  et  $\gamma(k_2, \lambda)$  est une loi  $\gamma(k_1 + k_2, \lambda)$ . On rappelle par ailleurs qu'une loi exponentielle est une loi gamma de paramètre  $\gamma(1, \lambda)$  et que la densité d'une loi gamma  $\gamma(k, \lambda)$  s'écrit  $f_{\gamma(k, \lambda)}(x) = \frac{(\lambda x)^{k-1}}{\Gamma(k)} \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}^+}$  où  $\Gamma$  désigne la fonction gamma.

$N([0, t])$  est donc bien fini p.s. et suit une loi de Poisson de paramètre  $\lambda t$ .  $[0, t]$  étant un élément générateur de la tribu borélienne sur  $\mathbb{R}$ , on en déduit le résultat pour  $N(A)$ .

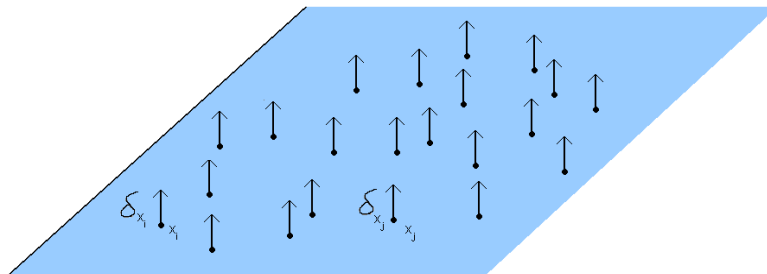
**Remarque 1.** Plus généralement, pour toute suite de points  $\left(S_n = \sum_{k=1}^n X_k\right)_{n \in \mathbb{N}^*}$  dont le processus de comptage est un processus à accroissements indépendants stationnaires (P.A.I.S.) - voir [7] - on peut voir que les accroissements  $X_k$  qui séparent chacun des points  $S_n$  sont indépendants et suivent une loi exponentielle de paramètre  $\lambda$ . On en déduit ensuite le résultat précédant concernant la loi du processus de comptage.

On peut aussi définir un processus ponctuel comme une variable aléatoire à valeur dans l'espace mesurable des configurations localement finie  $(\mathcal{N}^{lf}, \mathcal{N}^{lf})$ , où  $\mathcal{N}^{lf}$  est la plus petite tribu qui rend pour tout borélien  $A$  l'application  $\mathbf{x} \mapsto N(A)_{\mathbf{x}}$  mesurable.

**Définition 3 (Loi d'un processus ponctuel).** On appelle loi du processus ponctuel la mesure de probabilité  $\pi$  induite sur  $\mathcal{N}^{lf}$ .

**Remarque 2.** Il existe d'autres définitions de processus ponctuels. Une des plus répandue est celle qui définit un processus ponctuel en termes de mesures aléatoires, associant des masses aux points du processus (voir Figure 2).  $\mathcal{N}^{lf}$  représente alors l'ensemble des mesures de comptage aléatoires, réparties de manière localement finie.

Le processus ponctuel s'écrit dans ce cas :  $X = \sum_{i=1}^N \delta_{X_i}$ , où désigne  $\delta_{X_i}$  la mesure de Dirac au point  $X_i \in \mathcal{X}$ .



$(\chi, d)$

Figure 4

Il existe aussi une formulation en termes d'ensembles aléatoires, dont une est liée à la famille  $\{N(A)\}_A$  qui compte le nombre de points dans les boreliens  $A$ . Cette définition constructive requiert néanmoins la condition suivante : lorsque  $A \cap B = \emptyset$ ,  $N(A \cup B) = N(A) + N(B)$  presque sûrement.

De nombreuses références concernant les différentes manières de définir les processus ponctuels sont données dans [1].

### 1.3 Fidis

Cette courte partie définit un analogue des fonctions de répartitions pour les processus ponctuels.

**Définition 4 (Fidis).** Soit un  $X$  processus ponctuel sur  $(\chi, d)$ . La famille des lois marginales de dimension finie, appelé fidis (fidis pour Finite Dimensional Distributions) est l'ensemble des lois marginales  $(N(A_1), \dots, N(A_m))$ , quelque soit le vecteur de taille finie  $m$  considéré  $(A_1, \dots, A_m)$ , où  $A_1, \dots, A_m$  désignent des boreliens bornés quelconques de  $\chi$ .

L'analogie se fait ici :

**Théorème 1.** La loi d'un processus ponctuel sur un espace métrique complet séparable est entièrement déterminée par la donnée de la famille des fidis.

**Remarque 3.** Ce théorème implique que si deux processus ponctuels ont les même fidis, alors ils suivent la même loi.

▲ Esquisse de la preuve :

On remarque que les ensembles de la forme  $\{\omega \in \Omega : N_{X(\omega)}(A_i) \in B_i, i = 1 \dots m\}$  où les  $A_i$  désignent des boreliens bornés de  $\chi$  et  $B_i$  des boreliens de  $\mathbb{R}$  forment un semi-anneaux engendrant  $\mathcal{N}^{lf}$ . Par ailleurs, la probabilité de chacun de ces ensemble peut être déterminé à l'aide des fidis. On conclut en appliquant un théorème de classe monotone pour les semi-anneaux. Voir [3] □

### 1.4 Processus ponctuels marqués

On va maintenant associer aux points du processus ponctuels des marques. Cela s'avère assez utile en pratique lorsqu'on souhaite distinguer la nature des points : magnitude, spectre d'émission des étoiles, prédateurs ou proies, cellules marquées par des marqueurs radioactifs, espèces d'arbres présents dans une forêt, ...

**Définition 5 (Processus ponctuel marqué).** Soit  $(\chi, d)$  et  $(K, d')$  deux espaces métriques complets séparables. Un processus ponctuel marqué à positions dans  $\chi$  et à marques dans  $K$  est un processus ponctuel sur l'espace produit  $\chi \times K$  de sorte que le processus ponctuel non marqué soit bien définie.

Par processus non marqué, on entend processus obtenu par les projections des points de chaque marque sur une marque de référence.

**Exemple 3.** Sur la figure 5, on peut voir un exemple de processus marqué, processus ponctuel selon les marques  $K = \mathbb{R}$ , représenté par l'intermédiaire des différents plans, et processus ponctuel selon les points spatiaux de  $\chi = \mathbb{R}^2$ , représenté par les points sur chaque plan associé à une marque. Le processus non marqué résultant est obtenu par la projection des points spatiaux de chaque marque sur un plan de référence - plan défini par  $\chi \times \{0\}$  ici. Afin de ne pas alourdir le schéma, les projections des différents points marqués ne sont pas toutes explicitées.

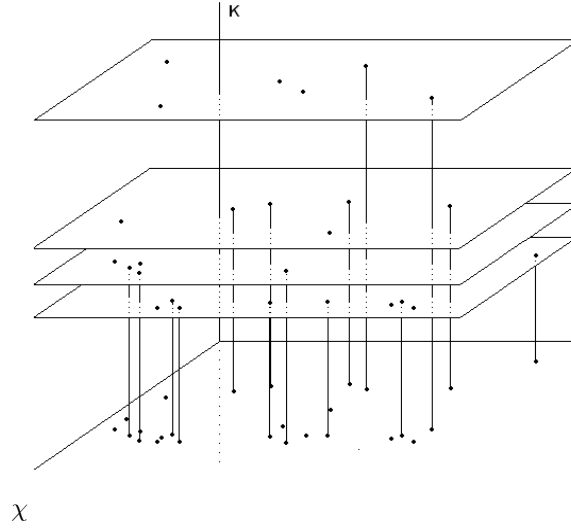


Figure 5

**Remarque 4.** Signalons ici que l'espace produit  $\chi \times K$  muni de la distance

$$\rho((x, k), (y, l)) = \max\{d((x, y), d'(k, l))\}$$

est séparable complet, ce qui place la définition précédente dans le cadre général fixé.

**Exemple 4.** Regardons le cas assez fréquent où l'ensemble des marques est de cardinal fini :  $K = \{1, \dots, M\}$ . On munit  $K$  de la distance suivante :  $d'(k, l) = |k - l|$ .

Tout processus ponctuel marqué  $Y$  défini sur le produit  $\chi \times K$  donne lieu à un processus ponctuel non marqué  $X$  bien défini.

En effet, soit  $A$  un borélien borné de  $\chi$ . Le nombre de points non marqués dans  $A$  peut s'écrire :

$$N_X(A) = \sum_{i=1}^M N_Y(A \times \{i\}) \quad (1)$$

où  $N_Y(A \times \{i\})$  désigne le nombre de points dans  $A$  ayant la marque  $i \in \{1, \dots, M\}$ . Or par définition,  $N_Y(A \times \{i\})$  est une variable aléatoire finie pour tout  $i \in \{1, \dots, M\}$ .  $N_X(A)$  est donc une variable aléatoire finie comme somme finie de variables aléatoires finies.

**Exemple 5.** Le processus ponctuel multi-varié  $Y = (Y_1, \dots, Y_M)$  peut être vue comme un processus ponctuel marqué dont les marques sont les indices des composantes. En effet, comme l'ensemble des marques est fini, le processus ponctuel non marqué  $X = \bigcup_{i=1}^M X_i$  est bien défini comme union finie de processus ponctuels.

**Remarque 5.** Dans le cas d'un ensemble de marques infini, dénombrable ou non, il faut être plus prudent.

On voit sur l'exemple 4 dans le cas dénombrable que si  $M$  tend vers  $\infty$ , la somme (1) peut diverger.

Pour un ensemble de marque ayant la puissance du continu (intervalle de  $\mathbb{R}$  par exemple), compter le nombre de point n'a pas vraiment de sens.

## 1.5 Processus ponctuels simples

La répartition des épïcètres peut présenter des points multiples dans le sens où deux épïcètres ou plus peuvent être situé exactement au même endroit. On parlera alors, au sein de la configuration, de *points multiples*.

Mais dans la plupart des exemples pratiques, une configuration de points ne peut avoir deux points ou plus au même endroit souvent parce que c'est physiquement impossible. Par exemple, il ne peut y avoir deux cellules distinctes ayant leurs centres confondus. On parlera alors, au sein de la configuration, de *points simples*.

Plus précisément, notons  $N_s^{lf}$  l'ensemble des configurations  $\mathbf{x}$  localement finies dont les points sont simples :  $N_{\mathbf{x}}(\{x\}) \in \{0, 1\}$  pour tout  $x \in \chi$ . Cet ensemble est  $\mathcal{N}^{lf}$ -mesurable car  $\chi$  étant séparable, il existe une suite dense  $(x_i)$  de sorte qu'on puisse recouvrir  $\chi$  par une union dénombrable de boules ouvertes  $B(x_i, 2^{-j})$  de rayons arbitrairement petits, et écrire  $N_s^{lf}$  sous la forme  $\bigcap_{i \geq 0} \{\omega \in \Omega : N(B(x_i, 2^{-j})) \in \{0, 1\}\} \in \mathcal{N}^{lf}$ .



**Définition 6.** Un processus ponctuel est dit simple s'il prend ses valeurs dans  $N_s^{lf}$  presque sûrement.

**Remarque 6 (définition - probabilités de vide).** Les processus ponctuels simples ont un avantage pratique, ils ne nécessitent pas de connaître une famille entière de fidis pour disposer de leur loi. En effet, il suffit seulement de connaître les probabilités de vide  $v(A) = \mathbb{P}(N(A) = 0)$  pour une collection suffisamment grande de borélien  $A$  de  $\chi$ . Pour plus de détails, on pourra consulter les références données en [1].

Le théorème qui suit sera admis mais le lecteur qui désire lire sa démonstration plutôt technique pourra la trouver dans [1] pages 10-12.

**Théorème 2.** La loi d'un processus simple  $X$  sur un espace séparable complet  $(\chi, d)$  est uniquement déterminé par les probabilités de vide  $v$  pour l'ensemble des boréliens bornés  $A \subseteq \chi$ . On peut aussi remplacer borélien par compact. Pour plus de détails, on pourra consulter les références données en [1].

**Exemple 6.** Le processus ponctuel vue en exemple 1 est un processus simple : il n'y a qu'un seul point. Sa probabilité de vide vaut par ailleurs :

$$v(A) = \mathbb{P}(U \in \chi \setminus A) = \frac{\mu(\chi \setminus A)}{\mu(\chi)}$$

**Exemple 7.** Le processus ponctuel de Poisson monodimensionnel vue en exemple 2 est aussi un processus simple car les accroissements  $X_k$  sont presque sûrement strictement positifs. Par ailleurs pour tout  $0 < s < t$  :

$$\begin{aligned} v([s, t]) &= \mathbb{P}(X_1 > t) + \sum_{k \geq 1} \mathbb{P}(X_1 + \dots + X_k \leq s; X_1 + \dots + X_k + X_{k+1} > t) \\ &= e^{-\lambda t} + \sum_{k \geq 1} \int_0^s \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x} e^{-\lambda(t-x)} dx \\ &= \sum_{k \geq 1} \frac{(\lambda s)^k}{k!} e^{-\lambda t} \\ &= e^{-\lambda(t-s)} \\ &= e^{-\lambda \mu([s, t])} \end{aligned}$$

**Remarque 7.** Deux processus  $X$  et  $Y$  quelconques et qui ont les mêmes probabilité de vide peuvent former deux processus simples  $X^s$  et  $Y^s$  de même probabilité de vide, en ignorant les points multiples. Par conséquent, ces derniers sont de même loi d'après le théorème précédent et les processus initiaux  $X$  et  $Y$  ont des lois qui ne diffèrent que dans la multiplicité des points.

## 1.6 Processus ponctuels finis

En plus d'être souvent simples, la plupart des configurations sont aussi finies. Cette finitude provient généralement de l'aspect borné de la fenêtre d'observation mais peut aussi provenir du fait que le processus génère simplement qu'un nombre fini de points.

### 1.6.1 Caractérisation d'un processus ponctuel fini

**Définition 7 (processus ponctuel fini).** On appelle processus ponctuel fini un processus dont les configurations sont finies.

Les processus ponctuels finis se modélisent bien par la donnée de :

- ▶ une loi de probabilité discrète  $(p_n)_{n \in \mathbb{N}}$ , qui régit le nombre de points.
- ▶ une famille de densité de probabilité  $j_n(x_1, \dots, x_n)$ ,  $n \in \mathbb{N}$ , symétriques en ses variables, qui régit la position des points sur  $\chi^n$ .

**Remarque 8.** On suppose ici que  $\chi$  est muni d'une mesure borélienne  $\nu(\cdot)$  de sorte que les densités  $j_n(x_1, \dots, x_n)$  soient définies par rapport à la mesure produit  $\nu(\cdot)^n$ .

Par ailleurs, l'aspect symétrique requis pour  $j_n$  provient du fait que les points générés par le processus ponctuels sont indifférents à l'ordre dans lequel on les indice. La densité  $j_n$  doit donc être la même quelque soit la manière d'indicer les points de la configuration. Pour plus de détails, on peut consulter [1] pages 13-14

### 1.6.2 Exemple fondamental : le processus binomial

Dans cette partie,  $\chi$  désigne un compact de  $\mathbb{R}^d$  de volume  $\mu(\chi) > 0$ .

**Définition 8 (Processus ponctuel binomial).** On appelle processus ponctuel binomial un processus ponctuel simple  $X = \{X_1, \dots, X_n\}$  constitué d'une union de  $n$  (fixé) points indépendants, répartis uniformément,  $X_1, \dots, X_n$ .

Comme  $P(N(\chi) = n) = 1$ , on a en particulier  $p_m = \delta_n^m$  et  $j_n(x_1, \dots, x_n) = \left(\frac{1}{\mu(\chi)}\right)^n$  du fait de l'uniforme distribution des points  $X_i$ .

Ce processus tire son nom du fait que pour tout borélien  $A \subseteq \chi$ ,

$$N(A) = \sum_{i=1}^n \mathbf{1}\{X_i \in A\}$$

suit une loi binomiale  $B(n, \frac{\mu(A)}{\mu(\chi)})$  qu'on note dans la suite  $P^{(n)}(\cdot)$ . En effet,  $N(A)$  compte le nombre de manières possibles de mettre  $n$  points dans  $A$  avec la probabilité  $\frac{\mu(A)}{\mu(\chi)}$ .

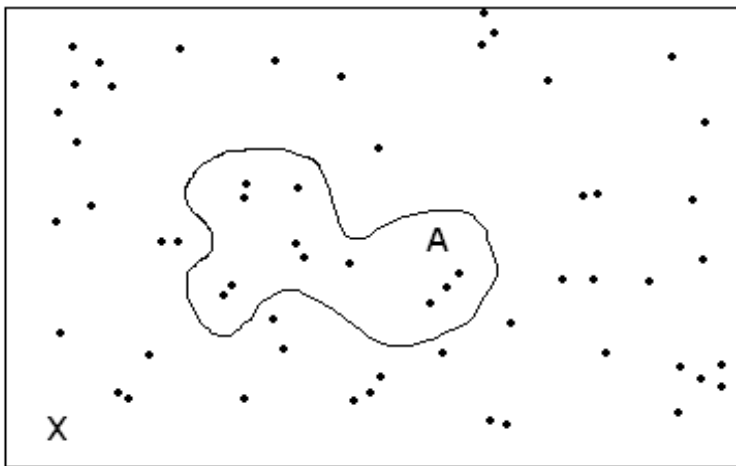


Figure 4.

Dans le cas où le compact  $\chi$  serait de mesure  $\mu(\chi)$  nulle, il est possible d'identifier  $\chi$  comme un sous espace de  $\mathbb{R}^c$  avec  $c < d$  dans lequel  $\mu(\chi) > 0$  :

**Exemple 8.** Le cercle unité est de mesure nulle dans  $\mathbb{R}^2$  mais peut être vue comme un ensemble de mesure non nulle  $\mu([0, 2\pi[)$  dans  $\mathbb{R}$ . Ainsi, si  $X_1, \dots, X_n$  sont des points répartis uniformément sur  $[0, 2\pi[$ , alors  $X = \{(\cos X_1, \sin X_1), \dots, (\cos X_n, \sin X_n)\}$  est un processus ponctuel binomial sur le cercle unité  $\chi \subset \mathbb{R}^2$ .

## 2 Processus de Poisson

Ces processus ponctuels sont sans doute les plus connus et les plus faciles à mettre en oeuvre. Ces processus ont la particularité de très bien représenter l'aléatoire spatial, notion que l'on précisera ensuite mais qui peut déjà bien s'envisager lorsqu'on pense par exemple aux points formés par les gouttes de pluie sur le macadam.

D'un point de vue pratique, lorsqu'on dispose d'une configuration de points, on teste d'abord son caractère aléatoire afin de voir si les points sont disposés selon une certaine logique (dépendance, différence de concentration, ...) ou non.

Il existe plusieurs manières de présenter le processus de Poisson. Dans ce qui suit, on présente une suite de résultats qui amène intuitivement à une définition rigoureuse du processus ponctuel de Poisson, à partir des processus binomiaux.

### 2.1 Processus de Poisson

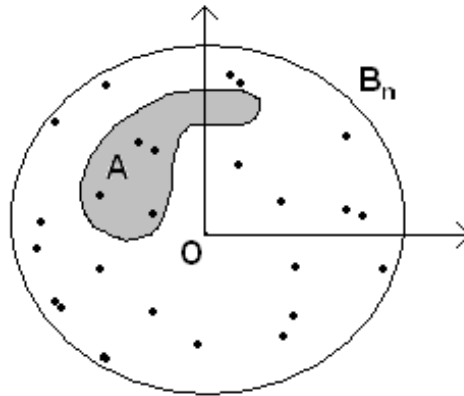
Avant d'aborder le processus de Poisson général, on définit d'abord certains cas particuliers de processus de Poisson : les processus de Poisson homogènes et inhomogènes.

#### 2.1.1 Processus de Poisson homogène

Le processus ponctuel de Poisson homogène présenté ici n'est qu'une généralisation du célèbre processus ponctuel de Poisson à une dimension.

Soit  $B_n \subset \mathbb{R}^d$  une boule centrée à l'origine, avec un rayon tel que son volume  $\mu(B_n)$  soit égal à  $\frac{n}{\lambda}$ , où  $\lambda$  désigne un réel strictement positif et  $P^{(n)}$  la loi d'un processus binomial à  $n$  points dans  $B_n$ . Soit  $A$  un borélien borné de  $\chi$  et  $k \in \mathbb{N}$ . Comme  $(B_n)_n$  est une suite croissante pour l'inclusion, dont l'union recouvre l'espace  $\chi$  tout entier, on peut trouver un entier  $n_0 \geq k$  tel que pour tout  $n \geq n_0$ , la boule  $B_n$  contient  $A$ . On a alors :

$$\begin{aligned} P^{(n)}(N(A) = k) &= P^{(n)}(N(A) = k ; N(B_n \setminus A) = n - k) \\ &= \binom{n}{k} \left( \frac{\mu(A)}{\mu(B_n)} \right)^k \left( \frac{\mu(B_n \setminus A)}{\mu(B_n)} \right)^{n-k} \end{aligned}$$



$$n = N(B_n) = N(A) + N(B_n \setminus A) = k + (n - k)$$

Figure 5.

Or, on sait que pour une loi binomiale quelconque  $\mathcal{B}(n, p)$ , on a :

$$\begin{aligned} \mathcal{B}(n, p) &\xrightarrow[n \rightarrow +\infty]{n.p \rightarrow \lambda} \mathcal{P}(\lambda) \end{aligned}$$

où  $\mathcal{P}(\lambda)$  désigne la loi de Poisson de paramètre  $\lambda$ .

Dans notre cas, étant donnée que  $n \frac{\mu(A)}{\mu(B_n)} = n \frac{\lambda \cdot \mu(A)}{n} = \lambda \cdot \mu(A)$  pour tout  $n \in \mathbb{N}$ , on a :

$$\mathcal{B}(n, \frac{\mu(A)}{\mu(\chi)}) \xrightarrow{n \rightarrow +\infty} \mathcal{P}(\lambda \cdot \mu(A)).$$

Si on choisit des boréliens bornés disjoints  $A$  et  $B$  inclus dans  $\chi$ , on peut de la même façon trouver un entier  $n_0$  de sorte que  $A \cup B$  soit contenu dans un n'importe quelle boule  $B_n$ , où  $n \geq n_0$ .

Soient  $k$  et  $l$  deux entiers naturels et  $n \geq \max(n_0, k + l)$ . On a, de manière analogue à ce qui précède :

$$\begin{aligned} P^{(n)}(N(A) = k ; N(B) = l) &= P^{(n)}(N(A) = k ; N(B) = l ; N(B_n \setminus A \cup B) = n - k - l) \\ &= \binom{n}{k} \left( \frac{\mu(A)}{\mu(B_n)} \right)^k \binom{n-k}{l} \left( \frac{\mu(B)}{\mu(B_n)} \right)^l \left( \frac{\mu(B_n \setminus A \cup B)}{\mu(B_n)} \right)^{n-k-l} \\ &\xrightarrow{n \rightarrow \infty} \frac{(\lambda \cdot \mu(A))^k}{k!} e^{-\lambda \mu(A)} \cdot \frac{(\lambda \cdot \mu(A))^l}{l!} e^{-\lambda \mu(A)} \cdot e^{-\lambda \mu(A \cup B)} e^{\lambda \mu(A)} e^{\lambda \mu(B)} \end{aligned}$$

On a donc :

$$\begin{aligned} \lim_{n \rightarrow \infty} P^{(n)}(N(A) = k ; N(B) = l) &= \frac{(\lambda \cdot \mu(A))^k}{k!} e^{-\lambda \mu(A)} \cdot \frac{(\lambda \cdot \mu(A))^l}{l!} \\ &= \lim_{n \rightarrow \infty} P^{(n)}(N(A) = k) \cdot \lim_{n \rightarrow \infty} P^{(n)}(N(B) = l) \end{aligned}$$

Par conséquent, lorsque  $A$  et  $B$  sont disjoints, les variables aléatoires  $N(A)$  et  $N(B)$  sont lorsque  $n \rightarrow \infty$  indépendantes et suivent par ailleurs une loi de Poisson.

Ces résultats suggèrent donc la définition suivante :

**Définition 9 (Processus Ponctuel de Poisson Homogène).** *Un processus ponctuel  $X$  sur  $\chi = \mathbb{R}^d$  est appelé processus ponctuel de Poisson homogène d'intensité  $\lambda > 0$  si :*

- $\forall A$  borélien borné,  $N(A)$  suit une loi de Poisson de paramètre  $\lambda \mu(A)$
- $\forall A_1, \dots, A_k$  boréliens bornés disjoints,  $N(A_1), \dots, N(A_k)$  sont indépendantes

**Remarque 9.** *La définition précédente vient dans un sens généraliser la notion de variable aléatoire réelles. En effet, si  $N$  désigne une variable aléatoire suivant une loi de Poisson, elle peut être vue comme le nombre moyen de points qui tombent dans  $[0, t[$  qui vaut par définition  $N([0, t]) = N_t = \lambda \cdot t = \lambda \cdot \mu([0, t])$ . Cette variable aléatoire n'est en fait rien d'autre qu'un processus ponctuel dans l'espace conteneur  $\chi = \mathbb{R}$ . On rappelle par ailleurs que l'esperance d'une telle variable est  $\lambda$ , ce qui peut justifier l'appellation intensité dans la définition précédente.*

**Remarque 10.** *Le coefficient  $\lambda$  est aussi appelé densité ou taux. Le terme densité permet parfois d'interpréter le coefficient  $\lambda$  autrement. Nous allons préciser cette interprétation dans ce qui suit.*

### 2.1.2 Processus ponctuel de Poisson inhomogène

Le coefficient  $\lambda$  peut être interprété comme une densité si l'on décide de rendre ce coefficient variable :  $\lambda = \lambda(x)$ ,  $x \in \chi$ . L'intensité devient variable selon l'endroit que l'on considère dans  $\chi$ . Donnons une définition pour préciser cette abstraction :

**Définition 10 (Processus Ponctuel de Poisson Inhomogène).** *On appelle processus ponctuel de Poisson inhomogène un processus ponctuel vérifiant :*

- $\forall A$  borélien borné,  $N(A)$  suit une loi de Poisson de paramètre  $\nu(A)$ ,  
où  $\nu(A) = \int_A \lambda d\mu$  et  $\lambda$  est une fonction positive, Lebesgue mesurable sur  $\chi$ .
- $\forall A_1, \dots, A_k$  boréliens bornés disjoints,  $N(A_1), \dots, N(A_k)$  sont indépendantes

La fonction  $\lambda$  peut donc être vue comme la dérivée de Radon Nikodym de  $\nu$  par rapport à la mesure de Lebesgue  $\mu$ .

### 2.1.3 Processus ponctuel de Poisson

On peut continuer la généralisation précédente en considérant des paramètres  $\nu$  sans densité. Donnons la définition d'un processus ponctuel de Poisson :

**Définition 11 (Processus Ponctuel de Poisson).** Soit  $\nu(\cdot)$  une mesure de Borel sur  $(\chi, d)$  espace métrique complet séparable, vérifiant  $\nu(\chi) > 0$  et  $\nu(A) < \infty$  pour tout borélien  $A$  borné de  $\chi$ . Une telle mesure est dite localement finie.

Un processus ponctuel sur  $\chi$  est un processus ponctuel de Poisson de mesure intensité  $\nu$  si

$$\blacktriangleright \forall A \text{ borélien borné, } N(A) \text{ suit une loi de Poisson de paramètre } \nu(A) \quad (P1)$$

$$\blacktriangleright \forall A_1, \dots, A_k \text{ boréliens bornés disjoints, } N(A_1), \dots, N(A_k) \text{ sont indépendantes} \quad (P2)$$

On voit à partir du cheminement précédent, consistant à partir d'une notion simple et visuelle pour aboutir à une définition assez générale et abstraite, la notion de densité émerger. En effet, les processus ponctuels de Poisson homogènes et inhomogènes sont des cas particuliers de processus ponctuel de Poisson où la mesure  $\nu$  admet une densité  $\lambda$  par rapport à la mesure de Lebesgue  $\mu$ .

**Remarque 11.** L'assertion (P2) peut s'interpréter comme la propriété d'aléatoire spatiale : des configurations de points situées dans des régions disjointes sont indépendantes entre elles. D'un point de vue pratique, cela implique d'une part qu'on puisse observer un processus de Poisson dans une fenêtre d'observation quelconque sans introduire aucun biais dans les raisonnements faits ensuite. D'autre part, les points qui sont en dehors de la fenêtre d'observation n'influent en rien les points observés.

Le théorème qui suit boucle le cheminement que nous avons suivi pour arriver aux processus de Poisson.

**Théorème 3.** Soit  $X$  un processus de Poisson sur un espace métrique séparable complet  $(\chi, d)$  de mesure intensité  $\nu(\cdot)$  et  $A$  un borélien borné de  $\chi$ . Alors, conditionnellement à l'évènement  $\{N(A) = n\}$ ,  $X$  restreint à  $A$  suit une loi binomial dont les points sont indépendants et disposés de manière  $\nu$ -uniforme.

▲Preuve : Soit  $B$  un borélien inclus dans  $A$ . La probabilité de vide de  $B$  conditionnellement à l'évènement  $\{N(A) = n\}$  s'écrit :

$$v(B|N(A) = n) = \mathbb{P}(N(B) = 0|N(A) = n) = \frac{\mathbb{P}(N(B) = 0 ; N(A \setminus B) = n)}{\mathbb{P}(N(A) = n)}$$

$X$  étant un processus ponctuel de Poisson,  $N(B)$  et  $N(A \setminus B)$  sont indépendants (P2) et suivent des lois de Poisson (P1). On en déduit donc :

$$v(B|N(A) = n) = \left( \frac{v(A \setminus B)}{v(A)} \right)^n$$

qui coïncide avec la probabilité de vide  $v(B)$  de  $n$  processus ponctuels indépendant et  $\nu$ -uniforme.  $B$  étant choisi de manière arbitraire, on en déduit le résultat avec le théorème 2.□

**Remarque 12.** Ce théorème s'avère en pratique très utile pour générer des processus de Poisson, car il permet de fixer (ou générer via une loi de probabilité) le nombre de points voulu pour une simulation. Cela n'est toutefois pas la seule façon de faire. M.N.M. van Lieshout propose, page 20, une référence sur ces problèmes.

**Remarque 13.** On peut rencontrer dans la littérature, et nous l'emploierons parfois ici, le terme de mesure aléatoire de Poisson.

### 2.1.4 Commentaires

Commençons par illustrer visuellement la différence entre un processus ponctuel homogène et inhomogène :



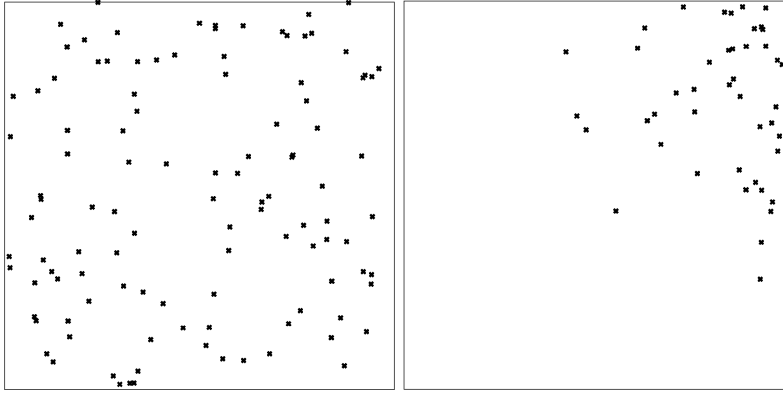


Figure 8 : Processus homogène (gauche) et processus inhomogène (droite) - [4]

Le processus inhomogène permet donc de modéliser des gradients ou des répartitions de points aléatoires en fonction d'un terrain que l'on peut paramétrer via la densité  $\lambda$ . On pourra aussi trouver d'autres illustrations dans [1] page 17, avec les paramètres de simulation page 19. Le théorème précédent peut être utilisé pour générer ces deux figures.

Notons toutefois que les processus de Poisson sont utilisés et très pertinents lorsqu'on n'impose aucune hypothèse sur la structure spatiale du phénomène étudié.

**Exemple 9 (Simuler un processus ponctuel de Poisson dans une zone "compliquée").** *Il peut arriver en pratique que la zone dans laquelle on étudie un phénomène soit assez accidentée ou peu sympathique sur le plan géométrique - une fractale par exemple. Pour simuler un processus de Poisson dans une telle zone, on le simule d'abord dans un compact de forme basique, un pavé mettons, qui contient la zone d'intérêt. On obtient alors le processus voulu en enlevant les points qui ne sont pas dans la zone souhaitée. Cette méthode porte le nom de rejection sampling et s'appuie sur la propriété (P2) du processus de Poisson.*

**Proposition 1 (Superposition).** *Soit  $X$  et  $Y$  deux processus ponctuels de Poisson indépendants sur  $(\chi, d)$  complet séparable de mesure intensité  $\lambda(\cdot)$  et  $\nu(\cdot)$  sans masse ponctuelle. Alors la superposition  $X \cup Y$  est un processus ponctuel de Poisson de mesure intensité  $\lambda(\cdot) + \nu(\cdot)$ .*

▲Preuve : La probabilité pour qu'un borélien  $A \subseteq \chi$  ne contienne aucun point vaut, en utilisant l'indépendance des processus :  $\mathbb{P}(X \cap A = 0; Y \cap A = 0) = e^{-\lambda(A)}e^{-\nu(A)} = e^{-[\lambda(A)+\nu(A)]}$ . Avec le théorème 2, on en déduit le résultat. □

**Remarque 14.** *Le théorème 3 donne un moyen de construire le processus de Poisson, ce qui fournit donc une preuve de son existence. Succinctement, on considère sur une partie bornée  $C \subset \chi$  un nombre  $n$  de points qui suit une loi de Poisson de paramètre  $\lambda\mu(C)$  et on suppose qu'il existe une suite de points indépendants  $(X_i)_i$  répartis de manière uniforme dans cette partie bornée. En posant  $X = \bigcup_{i \geq 0} X_i$ , conditionnellement à  $\{N(C) = n\}$ , on peut montrer que ce processus vérifie les deux assertions (P1) – (P2). Enfin, on peut construire un processus d'intensité  $\nu(\cdot)$  sur  $\chi$  tout en entier en partitionnant  $\chi$  en des parties bornées  $C_i$  de mesure  $\nu(C_i) > 0$  et en définissant sur chacune de ces parties, un processus  $X^{(i)}$  comme précédemment. On fait ensuite l'union de tout ces processus  $X^{(i)}$ .*

**Remarque 15.** *Le théorème 3 est aussi utile pour effectuer des simulations - voir [1] page 23.*

Nous allons maintenant aborder les processus définis à l'aide d'une densité par rapport à un processus de Poisson.

### 3 Processus fini défini à l'aide d'une densité

#### 3.1 Exemples introductifs : le processus agrégé et le modèle booléen

Dans de nombreuses applications ou expériences, il n'est pas réaliste de supposer que les points sont répartis de manière totalement aléatoire. Si les processus de Poisson ont l'avantage de répartir aléatoirement les points,

ils ont surtout la puissance de générer un grand nombre de modèles plus élaborés et parfois correspondant à un phénomène bien précis.

On peut à titre d'exemple citer les processus ponctuels de Poisson agrégés (Poisson cluster process) qui furent initialement introduits pour modéliser les agrégations de galaxies, en astronomie.

### 3.1.1 Le processus de Poisson agrégé : modèle de Neymann-Scott

Ce processus est défini de la manière suivante : chaque point "père" d'un processus ponctuel de Poisson donne naissance à un processus ponctuel fini de Poisson "fils" indépendant des ses "frères". On appelle l'union des points "fils" processus ponctuel de Poisson agrégé.

Lorsque les processus fils suivent tous la même loi, on appelle ce processus le modèle Neymann-Scott.

Donnons un autre exemple de processus qui fait partie de cette classe de modèle : le modèle de Matérn. On l'obtient en réalisant un processus de Poisson homogène d'intensité  $\lambda$  pour les points pères, autour desquels - dans un boule centrée sur un point père et de rayon  $r > 0$  - on place uniformément un nombre de points fils suivant un loi de Poisson de paramètre  $\nu$ . Le hasard intervient donc trois fois (positions des points pères, nombre de point fils, position des points fils). On peut bien évidemment varier les lois ce qui permet d'obtenir un grand nombre de modèles. Des références sont disponibles dans [1].

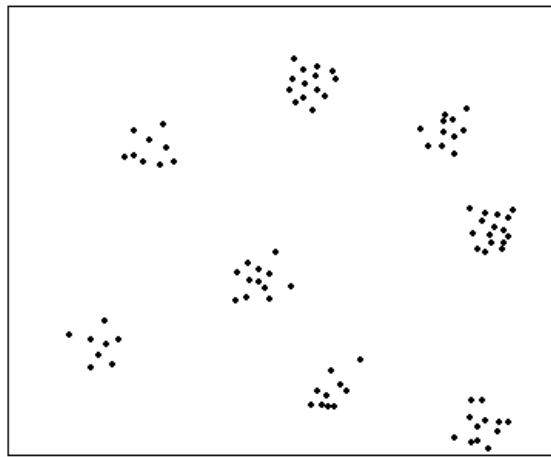


Figure 9 : Processus Matérn

### 3.1.2 Le modèle booléen

Le modèle booléen est lié au processus agrégé. On l'évoque pour l'intérêt qu'on lui a porté dans l'histoire...

Le modèle booléen est défini comme l'union des boules de rayon de  $r > 0$  fixé - on peut aussi mettre une loi sur  $r$  - centrées en des points d'un processus ponctuel de Poisson homogène.

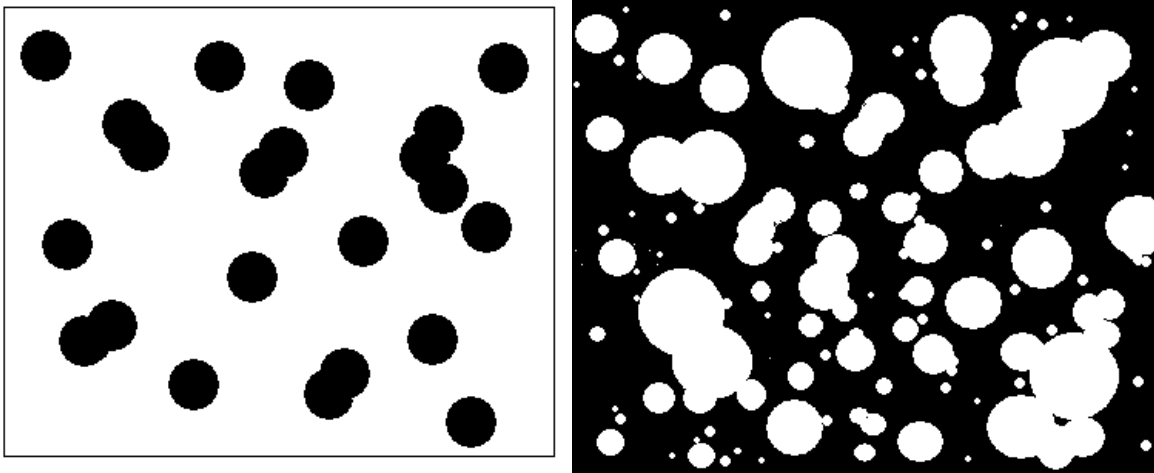


Figure 10 : le modèle booléen avec un rayon fixe et le bombing model avec un rayon de loi exponentielle

Ce modèle fut historiquement assez étudié pour les interprétations militaires qu'on pouvait lui donner. On peut en effet voir une boule du processus comme la zone dévastée par un obus qui serait tombé en un point père. Ce modèle est aussi connu sous le nom de *Bombing Model*. Voir [1] page 25 ainsi que les références données.

### 3.2 Définition et exemples

Nous allons voir maintenant sur le plan mathématique en quoi le processus de Poisson permet de modéliser un grand nombre de processus ponctuel finis. Nous allons modéliser un processus ponctuel à l'aide d'une densité de probabilité  $p$  conditionnellement à un processus de Poisson.

Cette approche n'est utile que pour les processus ponctuels finis car comme nous allons le voir avec la proposition qui suit, l'absolue continuité pour des processus ponctuels non fini est une hypothèse très forte. Par ailleurs, étant donnée une suite de densités sur une suite croissante d'ensembles de  $\mathbb{R}^p$  - par exemple - la limite n'existe pas toujours et si elle existe, n'est pas forcément unique. Pour plus de détails, on pourra consulter les références proposées par [1].

**Proposition 2.** *Soit  $X_\mu$  et  $X_\lambda$  deux processus de Poisson homogènes d'intensité  $\mu$  et  $\lambda$  respectivement, à valeur dans  $\mathbb{R}^p$  définis sur un même espace de probabilité  $(\Omega, A, \mathbb{P})$ . Si  $\mu$  et  $\lambda$  sont différents, alors la loi de  $X_\lambda$  n'est pas absolument continue par rapport à celle de  $X_\mu$ .*

▲Preuve : □

#### 3.2.1 Définition

On notera dans ce qui suit  $\pi(\cdot)$  la loi d'un processus ponctuel de Poisson sur  $\chi$  dont la mesure intensité  $\nu(\cdot)$  est finie, sans masse ponctuelle.

**Définition 12 (Densité de probabilité d'un processus ponctuel fini).** *Soit  $p : N^f \rightarrow \mathbb{R}^+$  une application mesurable positive vérifiant :*

$$\int_{N^f} p(\mathbf{x}) d\pi(\mathbf{x}) = 1.$$

Alors  $p$  est une densité de probabilité et définit un processus ponctuel  $X$  sur  $\chi$ . Le processus de Poisson dominant - on dit aussi de référence - étant fini et simple, il en est de même pour le processus  $X$ . Par ailleurs les éléments qui caractérisent  $X$  ( $(p_n)_{n \in \mathbb{N}}$ ,  $j_n(x_1, \dots, x_n)$ ) sont donnés en fonction de  $p$  par :

$$p_n = \frac{e^{-\nu(\chi)}}{n!} \int_{\chi^n} p(\{x_1, \dots, x_n\}) d\nu(x_1) \dots d\nu(x_n) \quad (3.1)$$

$$j_n(x_1, \dots, x_n) = \frac{p(\{x_1, \dots, x_n\})}{\int_{\chi^n} p(\{x_1, \dots, x_n\}) d\nu(x_1) \dots d\nu(x_n)} \quad (3.2)$$

Ces deux dernières formules peuvent se montrer mais nous les prendrons ici comme définition

#### 3.2.2 Exemples

**Processus de Poisson homogène d'intensité  $\lambda$**  Soit  $\chi$  un compact de  $\mathbb{R}^p$  et  $X$  un processus ponctuel Poisson homogène d'intensité  $\lambda$ . La densité de  $X$  est donnée par :

$$p(\{x_1, \dots, x_n\}) = \lambda^n e^{(1-\lambda)\mu(\chi)}.$$

#### Processus de Matèrn

**Processus de Gibbs** Les processus ponctuels de Gibbs viennent de la physique statistique, et caractérisent l'état d'équilibre des particules à l'intérieur de systèmes physiques fermés.

La densité d'un processus de Gibbs est donnée par

$$p(\{x_1, \dots, x_n\}) = \frac{e^{-U(\{x_1, \dots, x_n\})}}{Z}$$

où  $Z$  est une fonction de partition normalisante et  $U$  une fonction d'énergie, souvent dérivé d'un potentiel.

**Processus de Strauss**

## 4 Processus ponctuel de Markov

Jusqu'à maintenant, nous avons rencontré des processus ponctuels qui fournissent des configurations où les points présentent souvent une indépendance les uns des autres. En effet, pour le processus binomiale, tous les points sont indépendants. Pour le processus de Poisson, les nombres de points dans des régions disjointes sont indépendants, de même pour les modèles de processus à agrégat ou booléen, qui sont sous-jaçant à un processus de Poisson.

Dans ce chapitre, on présente des processus ponctuels - les processus ponctuels de Markov - qui peuvent prendre en compte des interactions présentes entre les points et qui incluent aussi des propriétés analogues aux processus de Poisson.

### 4.1 Exemple introductif et rappels sur les chaînes de Markov

#### 4.1.1 Le modèle des sphères dures

Cet modèle s'adapte bien à une problématique rencontrée en biologie : modéliser la disposition des cellules.

On suppose ici que les cellules sont toutes du même type et forment approximativement des sphères de rayon  $R > 0$  fixe. Pour modéliser la disposition des cellules, il convient de modéliser celle des noyaux dont on suppose qu'ils sont le centre des cellules.

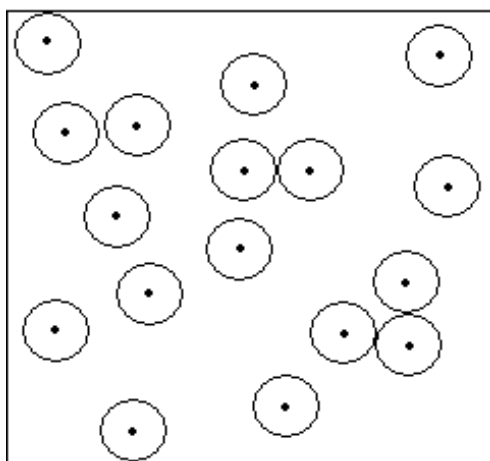


Figure 11 : Modèles des sphères dures

Il est clair que deux noyaux ne peuvent pas être plus proches que d'une distance  $2R$ . Un choix alors naturel s'impose pour la densité du processus :

$$p(\{x_1, \dots, x_n\}) = \alpha \beta^n \mathbf{1}(\|x_i - x_j\| > 2R, i \neq j)$$

où  $\beta$  désigne un paramètre d'intensité et  $\alpha$  une constante normalisatrice. Cette application est bien définie en tant que densité : elle est mesurable car  $\mathbf{1}(\|x_i - x_j\| > 2R, i \neq j)$  est borélienne et elle est symétrique en ses variables car  $\mathbf{1}(\|x_i - x_j\| > 2R, i \neq j)$  l'est. On peut voir un exemple de réalisation figure 11 où le contour des sphères a été rajouté.

### 4.1.2 Rappels sur les chaînes de Markov

On suppose connu les définitions de probabilité de transition et de chaîne de Markov. On pourra toutefois les revoir si nécessaire dans [2] chapitre 4. On rappelle ici juste la propriété - en général mieux connue que la définition - que possède une chaîne de Markov sur un espace dénombrable  $E$ .

Soit  $(X_i)_i$  une chaîne de Markov sur un espace  $E$  dénombrable, de matrice de transition  $(P_i)_i$ . On a alors pour tout  $x_1, \dots, x_{n-1}, x_n \in E$  :

$$\mathbb{P}(X_n = x_n \mid X_1 = x_1, \dots, X_{n-1} = x_{n-1}) = \mathbb{P}(X_n = x_n \mid X_{n-1} = x_{n-1}) = P_n(x_{n-1}, x_n) \quad (4.1)$$

Cette dernière égalité permet d'écrire la loi jointe sous forme de produit où n'intervient que la loi initiale et les probabilités de transition  $(P_i)_i$  :

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_0 = x_0) \prod_{i=1}^n P_i(x_{i-1}, x_i). \quad (4.2)$$

Ces égalités signifient que la connaissance du futur ne dépend que du présent - l'indice  $n$  s'interprète souvent comme un paramètre du temps. Néanmoins, dans notre cadre, il serait pratique de pouvoir considérer  $n$  comme un paramètre d'espace. Cela est assez difficile compte tenu de l'ordre imposé naturellement par le cours du temps. On pourrait essayer d'améliorer la formule (4.1), dans le cas de la dimension 1 en prenant :

$$\mathbb{P}(X_i = x_i \mid X_j = x_j ; j \neq i) = \mathbb{P}(X_i = x_i \mid X_{i-1} = x_{i-1} ; X_{i+1} = x_{i+1}) \quad (4.3)$$

et en adaptant pour les bords (0 et  $N$  si  $E$  est fini ou bornée dans un direction).

On peut aussi améliorer en dimension supérieure mais cela devient vite lourd au fur et à mesure que la dimension augmente car il faut regarder tout les points qui sont à une portée fixée du point sur lequel on se trouve. Schématiquement :

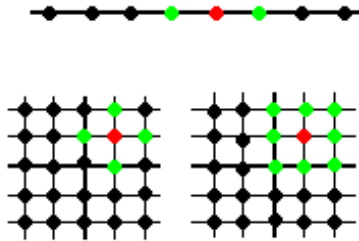


Figure 12 : Chaîne de Markov améliorée en dimension 1 et 2 - voisins (vert) du point rouge

Par ailleurs, il semble difficile d'avoir une formule de la forme (4.2) avec cette formulation.

## 4.2 Notions de voisinage

On introduit une notion de voisinage pour généraliser la notion de chaîne de Markov et l'étendre aux processus ponctuels dans un espace de dimension quelconque.

**Définition 13 (voisin).** Soit  $\sim$  une relation binaire symétrique et réflexive sur  $(\chi, d)$  - dite relation de voisinage. Soit  $u$  et  $v$  deux points de  $\chi$ . On dit que  $u$  et  $v$  sont voisins si  $u \sim v$ .

**Exemple 10 (La relation de portée fixe).** Soit  $R > 0$ . On définit la relation de portée  $R$  par :  $u \sim v \Leftrightarrow d(u, v) < R$ . En d'autres termes, les boules  $B(u, \frac{R}{2})$  et  $B(v, \frac{R}{2})$  sont d'intersection vide.

**Définition 14 (voisinage).** Soit  $A$  un ensemble  $\chi$ . On définit le voisinage de  $A$  par

$$\partial(A) = \{x \in \chi : x \sim a, a \in A\}$$

En particulier, le voisinage de  $\{a\}$  contient tous les voisins de  $a$ .



### 4.3 Densité conditionnelle de Papangelou

Dans cette partie, on n'expose uniquement les résultats et le sens qu'on peut leur prêter. Pour connaître les bases théoriques sur lesquelles s'appuie cette partie (moment d'un processus ponctuel, formules de Campbell-Mecke, théorie de Palm, ...) on peut consulter [1] pages 35-42.

Nous allons formaliser la loi d'un processus ponctuel  $X$  en un point  $x \in \chi$  conditionnellement à sa configuration sur  $\chi \setminus \{x\}$  - conditionnement extérieur. Ce formalisme est donné par la densité conditionnelle de Papangelou. A titre informatif, le problème dual de connaître la loi d'un processus  $X$  sachant qu'il y a déjà un point en  $x \in \chi$  - conditionnement intérieur - se traite d'une manière analogue avec les lois de Palm.

#### 4.3.1 Interprétation de la densité conditionnelle de Papangelou

**Définition 15 (Densité conditionnelle de Papangelou).** La densité conditionnelle de Papangelou notée  $\lambda(.,.)$  est définie de manière heuristique par l'équation suivante :

$$\lambda(x, \mathbf{x})dx = \mathbb{P}(N(dx) = 1 | X \cap (dx)^c = \mathbf{x} \cap (dx)^c) \quad (4.4)$$

Cette équation s'interprète comme la probabilité qu'à le processus  $X$  d'envoyer un point dans une région  $dx$  autour d'un point  $x \in \chi$  qui respecte aussi la configuration déjà existante à l'extérieur de  $dx$ .

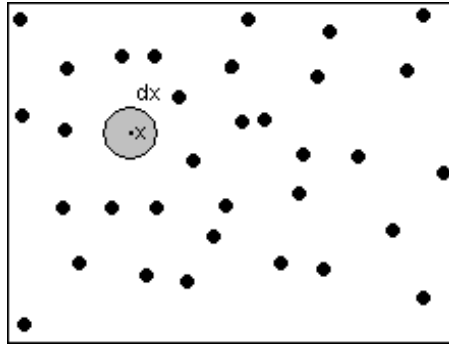


Figure 13 : interprétation de la densité conditionnelle de Papangelou

Pour définir de manière rigoureuse cette densité, il faut introduire la notion de mesure de Campbell réduite, que nous n'aborderons pas mais qu'on peut trouver dans [1] page 39.

#### 4.3.2 Une expression de la densité conditionnelle de Papangelou

Nous allons maintenant énoncé un théorème qui est parfois pris comme une définition de la densité conditionnelle de Papangelou :

**Théorème 4 (admis).** Soit  $X$  un processus ponctuel de densité  $p(.)$  par rapport à un processus de Poisson de mesure intensité  $\nu(.)$  finie, sans masse ponctuelle. Alors  $X$  a une densité conditionnelle de Papangelou

$$\lambda(u; \mathbf{x}) = \frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})} \quad (4.5)$$

pour tout  $u \notin \mathbf{x} \in N^f$ .

### 4.4 Processus ponctuel de Markov

On est maintenant en mesure de définir le processus ponctuel de Markov :

**Définition 16 (Processus ponctuel de Markov).** Soit  $(\chi, d)$  un espace métrique complet séparable,  $\nu(.)$  une mesure de Borel sans masse ponctuelle,  $\pi_\nu(.)$  la loi d'un processus de Poisson de mesure intensité  $\nu(.)$ . Soit  $X$  un processus ponctuel sur  $\chi$  de densité  $p(.)$  par rapport à  $\pi_\nu(.)$ . On dit  $X$  est un processus ponctuel de Markov par rapport à une relation de voisinage  $\sim$  si pour toute configuration  $\mathbf{x} \in N^f$  telle que  $p(\mathbf{x}) > 0$  :

$$\blacktriangleright p(\mathbf{y}) > 0 \text{ pour tout } \mathbf{y} \subseteq \mathbf{x}. \quad (a)$$

$$\blacktriangleright \forall u \in \chi, \frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})} \text{ ne dépend que de } u \text{ et } \partial(\{u\}) \cap \mathbf{x} = \{x \in \mathbf{x} : u \sim x\} \quad (b)$$

**Remarque 16.** Les hypothèses qui portent sur  $\nu(\cdot)$  font de  $X$  un processus fini et simple.

(a) signifie que si une configuration peut se produire, alors toutes les sous-configurations qu'elle contient peuvent se produire aussi. La condition (b) exprime une propriété de Markov locale : le comportement d'un point  $u$  par rapport à la configuration entière ne dépend que des ses proches voisins dans cette configuration. Enfin, si  $u \notin \mathbf{x}$  la condition (b) s'exprime à l'aide de la densité conditionnelle de Papangelou et fournit donc un équivalent spatial de la relation (4.3).

## 4.5 Théorème et propriétés fondamentales

### 4.5.1 La markovianité spatiale

Le théorème suivant reformule en terme de voisinage la définition précédente :

**Théorème 5.** Soit  $X$  un processus ponctuel de Markov de densité  $p(\cdot)$  et soit  $A$  un borélien de  $(\chi, d)$  espace métrique complet séparable. Alors la loi conditionnel de  $X \cap A$  sachant  $X \cap A^c$  ne dépend que de  $X$  restreint au voisinage  $\partial(A) \cap A^c$ .

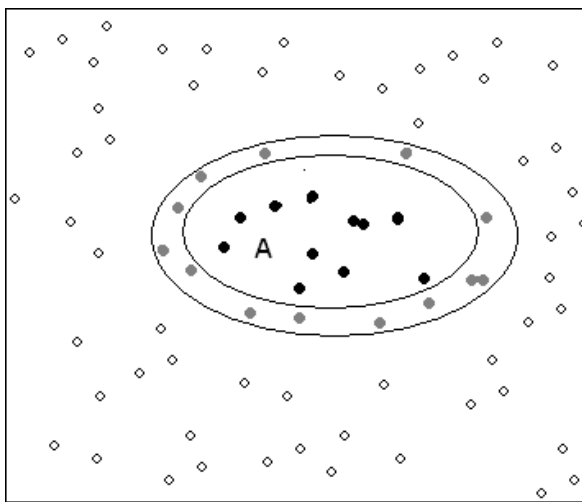


Figure 13 : processus markovien

▲Preuve : Soit  $\mathbf{z}$  une configuration. On partitionne  $\mathbf{z}$  selon les points dans  $A$  et dans  $A^c$  :  $\mathbf{z} = \mathbf{z}_{A^c} \cup \mathbf{z}_A$ .

...

□

On propose maintenant une écriture équivalente de l'équation (4.2) pour les processus ponctuels en introduisant la notion de clique :

### 4.5.2 Le théorème d'Hammersley - Clifford

**Définition 17 (Clique).** Soit  $\sim$  une relation de voisinage sur  $\chi$ . Une configuration  $\mathbf{x}$  est appelée clique si tous les points de  $\mathbf{x}$  sont voisins les uns des autres :  $\forall u, v \in \mathbf{x}, u \sim v$ . Par convention,  $\emptyset$  sera aussi une clique.

**Exemple 11.** Sur le schéma ci-dessous, la configuration possède 4 cliques non vides pour la relation de portée fixe  $R > 0$ .

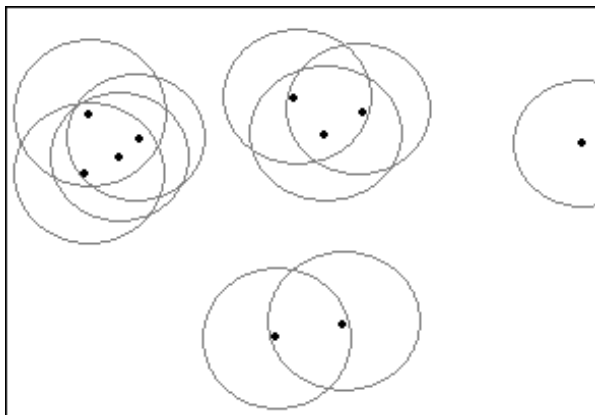


Figure 14 : configuration à 4 cliques

**Théorème 6 (Hammersley - Clifford).** *Une densité de processus ponctuel  $p : N^f \rightarrow [0, +\infty[$  est markovienne i.e. définit un processus ponctuel de Markov par rapport à la relation de voisinage  $\sim$  si et seulement si il existe une fonction  $\phi : N^f \rightarrow [0, +\infty[$  mesurable telle que*

$$p(\mathbf{x}) = \prod_{\text{clique } \mathbf{y} \subseteq \mathbf{x}} \phi(\mathbf{y})$$

pour tout  $\mathbf{x} \in N^f$ . On peut remplacer le produit sur les cliques par un produit sur toutes les sous-configurations  $\mathbf{y} \subseteq \mathbf{x}$  en prenant  $\phi(\mathbf{y}) = 1$  lorsque  $\mathbf{y}$  n'est pas une clique.

▲Preuve : Supposons que ....

...

□

## 4.6 Exemples

Donnons à présent des exemples de processus ponctuel de Markov.

Soit  $\chi$  un compact de  $\mathbb{R}^p$  de volume  $\mu(\chi) > 0$ .

### 4.6.1 Processus ponctuel de poisson d'intensité constante

Soit  $X$  un processus de Poisson homogène d'intensité  $\lambda$ . Alors  $X$  est markovien.

En effet, sa densité qui s'écrit pour toute configuration  $\mathbf{x} : p(\mathbf{x}) = e^{(1-\lambda)\mu(\chi)} \lambda^{\#\mathbf{x}}$  est toujours strictement positive, ce quelque soit la configuration  $\mathbf{x}$  considérée. Par ailleurs la densité conditionnelle de Papangelou

$$\lambda(u; \mathbf{x}) = \frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})} = \lambda \cdot \mathbf{1}\{u \notin \mathbf{x}\}.$$

Par conséquent,  $p(\cdot)$  est markovienne, et ce pour n'importe quelle relation de voisinage. C'est donc bien en accord avec le fait qu'un processus de Poisson homogène représente l'aléatoire spatial le plus total.

On peut tout aussi bien montrer ce point avec le théorème d'Hammersley-Clifford, en factorisant la densité à l'aide d'une fonction mesurable. On a :

$$p(\mathbf{x}) = e^{(1-\lambda)\mu(\chi)} \lambda^{\#\mathbf{x}} = e^{(1-\lambda)\mu(\chi)} \prod_{x \in \mathbf{x}} \lambda$$

où  $\phi(\emptyset) = e^{(1-\lambda)\mu(\chi)}$ ,  $\phi(\{u\}) = \lambda$  pour tout  $u \in \chi$  et  $\phi \equiv 1$  pour des configurations de cardinal supérieur ou égal à deux. Ceci confirme encore une fois le manque de dépendance entre les points.

### 4.6.2 Le modèle des sphères dures

Le processus associé au modèle des sphères dures développé en introduction de ce chapitre est un processus markovien de porté  $2R$ .

Remarquons déjà que sa densité  $p(\{x_1, \dots, x_n\}) = \alpha \beta^n \mathbf{1}(\|x_i - x_j\| > 2R, i \neq j)$  n'est strictement positive que pour des configurations  $\mathbf{x}$  dont les points sont tous à une distance d'au moins  $2R$ . Par conséquent, lorsque  $p(\mathbf{x}) > 0, \forall \mathbf{y} \subseteq \mathbf{x}, p(\mathbf{y}) > 0$  car tous les points de  $\mathbf{y}$  sont à une distance d'au moins  $2R$ .

Enfin, la densité conditionnelle de Papangelou en  $u$  sachant  $\mathbf{x} = \{x_1, \dots, x_n\}$  vaut :

$$\lambda(u; \mathbf{x}) = \beta \cdot \mathbf{1}(\|u - x_j\| > 2R, i \neq j) = \beta \cdot \mathbf{1}(\partial(\{u\}) \cap \mathbf{x} = \emptyset)$$

qui ne dépend donc que de  $u$  et de ses voisins dans  $\mathbf{x}$ . On en déduit que ce processus est markovien de porté  $2R$ .

On aurait pu le voir avec le théorème d'Hammersley-Clifford en posant :  $\phi(\emptyset) = \alpha, \phi(\{u\}) = \beta, \phi(\{u, v\}) = \mathbf{1}(\|u - v\| > 2R)$  et  $\phi \equiv 1$  pour des configurations de cardinal supérieur ou égal à trois

### 4.6.3 Cas général : le modèle des paires interactives

Le modèle des *paires interactives* (*pairwise interaction model*) est caractérisé par des densités de la forme

$$p(\mathbf{x}) = \alpha \prod_{x \in \mathbf{x}} \beta(x) \prod_{u, v \in \mathbf{x}; u \sim v} \gamma(u, v)$$

où  $\alpha$  est une constante positive normalisatrice,  $\beta : \chi \rightarrow [0, +\infty[$  une fonction d'intensité et  $\gamma : \chi \times \chi \rightarrow [0, +\infty[$  une fonction d'interaction entre les paires. Le modèle des sphères dures est donc un cas particulier où  $\beta$  est une fonction constante,  $\gamma(u, v) = \mathbf{1}(u \approx v)$ .

En prenant  $\gamma \equiv 1$ , on obtient par ailleurs un processus de Poisson homogène.

Le théorème d'Hammersley-Clifford assure que ce processus est markovien et comme on peut le voir avec ce modèle, ce théorème permet de définir de nouveaux modèles. Par ailleurs, au lieu de définir la densité  $p$  pour toute configuration  $\mathbf{x}$ , il suffit de définir les fonctions d'interactions  $\phi$  tout en faisant attention à ce que  $p$  reste une densité (intégrable en autre).

## 5 Processus Ponctuels Markoviens Marqués

Nous avons déjà présenté l'intérêt des processus marqués. Nous étendons simplement ces derniers aux processus markoviens.

### 5.1 Définitions

**Définition 18 (processus ponctuel markovien marqué).** Soient  $(\chi, d)$  et  $(K, d')$  deux espaces métriques complets séparables,  $\nu(\cdot)$  une mesure de Borel finie sur  $\chi$  sans masse ponctuelle,  $m(\cdot)$  une mesure de probabilité sur la tribu borélienne de  $K$  et  $\pi_{\nu \times m}(\cdot)$  la loi d'un processus de Poisson sur  $\chi \times K$  de mesure intensité :  $\nu \times m$ . Soit par ailleurs  $Y$  un processus ponctuel marqué à position dans  $\chi$  et à marque  $K$ , de densité  $p(\cdot)$  par rapport à la mesure aléatoire de Poisson  $\pi_{\nu \times m}(\cdot)$ . Alors  $Y$  est un processus ponctuel markovien marqué par rapport à la relation de voisinage  $\sim$  sur  $\chi \times K$  si pour toute configuration  $\mathbf{y}$  telles que  $p(\mathbf{y}) > 0$  :

$$\blacktriangleright p(\mathbf{z}) > 0 \quad \forall \mathbf{z} \subseteq \mathbf{y} \tag{I}$$

$$\blacktriangleright \forall (u, l) \in \chi \times K : \frac{p(\mathbf{y} \cup \{(u, l)\})}{p(\mathbf{y})} \text{ ne dépend uniquement que de } (u, l) \tag{II}$$

$$\text{et de } \partial(\{(u, l)\}) \cap \mathbf{y} = \{(x, k) \in \mathbf{y} : (u, l) \sim (x, k)\}$$

**Remarque 17.** Le théorème d'Hammersley-Clifford est toujours valable pour les processus ponctuel markovien marqués. Par conséquent, une densité  $p(\cdot)$  définit un processus ponctuel markovien marqué par rapport à la relation de voisinage  $\sim$  sur  $\chi \times K$  si et seulement si elle est peut être factorisée sous la forme

$$p(\mathbf{y}) = \prod_{\text{clique } \mathbf{z} \subseteq \mathbf{y}} \phi(\mathbf{z})$$

pour toute configuration finie  $\mathbf{y} \in N^f$ . ( $\phi$  fonction d'interaction).

## 5.2 Exemples

## 6 Conclusion

On voit que l'on peut modéliser sans trop de difficultés un grand nombre de situations courantes.

Par ailleurs, les modélisations vues ici peuvent s'avérer particulièrement utile pour faire des simulations et effectuer des pronostiques.

En lisant par exemple l'article [5], on constate que les images satellites ou aériennes fournissent des nombreuses données ponctuelles qui réagissent bien aux modèles mathématiques qui précèdent et ces derniers permettent de savoir, dans le cas d'une espèce de plante fragile quelles seront ses chances de survie, selon qu'elle constitue un configuration à points isolés, ou bien à points entourés de plantes compétitrices.

En ce qui concerne les processus ponctuels markoviens, ils sont très utiles pour faire de l'amélioration d'image entre autre, notamment parce que chaque pixel d'une image analysée par le filtre est considéré avec ces points voisins, ce qui permet lorsqu'on transfère des données sur un canal bruité de bien corriger les erreurs. C'est ce qu'on peut voir dans [4]. Ils sont aussi utiles pour la spectroscopie et l'étude des galaxies [6].

## Références

- [1] Markov Point Processes and their Applications, M. N. M. van LIESHOUT
- [2] Martingales et Chaîne de Markov, chapitre 4, BALDI, MAZLIAK, PRIOURET
- [3] Probability and Measure, P. BILLINGSLEY
- [4] Méthodes stochastiques en analyse d'image : des champs de Markov aux processus ponctuels marqués, X. DESCOMBES
- [5] Evaluation des Ressources Forestières à l'aide de Processus Ponctuels Marqués, G. PERRIN, X. DESCOMBES, J. ZEDUBIA
- [6] Restauration de processus ponctuels marqués pour la spectroscopie et l'astronomie, V. MAZET
- [7] Processus de Saut, J. BERTOIN
- [8] Processus ponctuels pour l'extraction de réseau linéique dans les images satellitaires et aériennes, R. STOICA