Processus de Poisson et meth. actuariels	M1 MMD
Gubinelli Massimiliano	Poly 1 - v.3 20120216

1 Le processus de Poisson

1 Processus de renouvellement et de comptage

Un *processus de renouvellement* est une suite croissante $(T_n)_{n\in\mathbb{N}}$ de v.a.s telles que les *temps inter-arrivée* $(\tau_n = T_n - T_{n-1} \geqslant 0)_{n\geqslant 1}$ forment une famille iid de v.a. positives. Un processus de renouvellement décrit les temps dans lesquels certains événements aléatoires se produisent: arrivée de clients à un magasin, appels téléphoniques à un standard, temps de replacement de materiel défectueux (des ampoules par exemple), arrivées des demandes de remboursement à un institut d'assurance, tremblements de terre, accidents de voiture, decadiments radioactives, etc...

Associée à un processus de renouvellement donné il y a le correspondent *processus de comptage* $(N_t)_{t\in\mathbb{R}_+}$ qui compte le nombre d'événements jusqu'à l'instant $t\in\mathbb{R}_+$

$$N_t = \#\{n \geqslant 0 : T_n \leqslant t\} = \sum_{n\geqslant 0} \mathbb{I}_{T_n\leqslant t}. \qquad (1)$$

Un processus de comptage est donc une famille de v.a. $(N_t)_{t\in\mathbb{R}_+}$ à valeurs dans $\mathbb{N}$ et indexé par la demi-droite réelle $\mathbb{R}_+ = [0, +\infty[$. C'est un exemple (avec le mouvement Brownien) de processus aléatoire à temps continu.

Remarque 1. La loi d'un processus à temps continu $(X_t)_{t\in I}$ est donnée par la collection des lois des tous les n-uplets $(X_{t_1}, \ldots, X_{t_n})$ pour tout $n \geqslant 1$ et toute choix $(t_i \in I)_{i=1,\ldots,n}$.

Définition 2. *Un processus continu $(X_t)_{t\in\mathbb{R}}$ est continu à droite ssi $X_{t+} = \lim_{s\to t, t\leqslant s} X_s = X_t$ pour tout $t\in\mathbb{R}$ presque sûrement. Il admet limite à gauche si $X_{t-} = \lim_{s\to t, s\leqslant t} X_s$ pour tout $t\in\mathbb{R}$ existe p.s. . Il est à accroissement indépendants si pour tout $n\geqslant 0$ et tout $t_1 \leqslant \cdots \leqslant t_n$ les v.a. $(X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$ sont indépendantes. Il est à accroissements stationnaires si pour tout $t \leqslant s$ et $h \geqslant 0$ la loi de $X_t - X_s$ est la même que la loi de $X_{t+h} - X_{s+h}$.*

Définition 3. *Un processus de comptage $(N_t)_{t\geqslant 0}$ est un processus à temps continu et à valeurs dans $\mathbb{N}\cup\{+\infty\}$, croissant, continu à droite (cad) avec $N_0 = 0$.*

Donné un processus de comptage on peut definir les temps aléatoires

$$T_n = \inf\{t \geqslant 0 : N_t \geqslant n\}$$

(avec la convention que $\inf \varnothing = +\infty$). C'est la suite de temps de saut de $(N_t)_{t\geqslant 0}$, elle est une suite croissante avec $T_0 = 0$, $T_1 > 0$ (consequence de la continuité à droite, exercice).

2 Le processus de Poisson

Si $(T_n)_{n\geqslant 0}$ est le processus de renouvellement avec des temps d'interarrivée de loi exponentielle de paramètre $\lambda > 0$ alors le processus de comptage correspondant est appelé un *processus de Poisson* (PP).

On rappelle les propriétés suivantes de la loi exponentielle:

1. La loi de la somme de n v.a. $\mathcal{E}(\lambda)$ indépendantes est la loi $\Gamma(n,\lambda)$. La densité de la loi Gamma $\Gamma(\alpha,\beta)$ de paramètres $\alpha>0,\beta>0$ est donnée par

$$f_{\Gamma(\alpha,\beta)}(x)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\,e^{-\beta x}\mathbb{I}_{x>0}\,.$$

2. Absence de mémoire: si $X\sim\mathcal{E}(\lambda)$ alors

$$\mathbb{P}(X>t+h|X>t)=\mathbb{P}(X>h)=e^{-\lambda h}$$

pour tout $t,h\geqslant 0$.

Théorème 4. *Si T est une v.a. positive telle que $\mathbb{P}(T>t+h|T>t)=\mathbb{P}(T>h)$ alors, il existe $\lambda>0$ tel que $T\sim\mathcal{E}(\lambda)$*

Démonstration. Soit $g(t)=\mathbb{P}(T>t)$. On a que $g(t+s)/g(t)=g(s)$. Donc $g(k)=g(1)^k$ pour tout $k\in\mathbb{N}$. $g(k)=g(k/p)^p$ et $g(k/p)=g(1)^{k/p}$. Par densité et continuité (à droite) de g: $g(x)=g(1)^x$ pour tout $x\geqslant 0$. On doit avoir $0<g(1)<1$ car si $g(1)=0$ alors $T=0$ et si $g(1)=1$ alors $T=+\infty$. Alors, avec $\lambda=-\log g(1)>0$ on a que $g(t)=e^{-\lambda t}$. □

On a que la loi de T_n est donnée par

$$f_{T_n}(t)=\frac{\lambda^{n-1}}{(n-1)!}t^{n-1}e^{-\lambda t}\mathbb{I}_{t>0}$$

et que

$$\mathbb{P}(T_n\geqslant t)=\int_0^{+\infty}\frac{\lambda^{n-1}}{(n-1)!}t^{n-1}e^{-\lambda t}\mathrm{d}t=\sum_{k=0}^{n-1}\frac{(\lambda\,t)^k}{k!}e^{-\lambda t}\qquad t\geqslant 0\,.$$

Mais alors il est facile de voir que

$$\mathbb{P}(N_t\geqslant n)=\mathbb{P}(T_n\leqslant t)=1-\sum_{k=0}^{n-1}\frac{(\lambda\,t)^k}{k!}e^{-\lambda t}$$

avec $\mathbb{P}(N_t\geqslant 0)=1$. Ce qu'implique que

$$\mathbb{P}(N_t=n)=\mathbb{P}(N_t\geqslant n)-\mathbb{P}(N_t\geqslant n+1)=\frac{(\lambda\,t)^n}{n!}e^{-\lambda t}$$

pour tout $n\geqslant 0$. La loi de N_t est donc une loi de Poisson de paramètre $\lambda\,t$ pour tout $t\geqslant 0$ (avec la convention qu'une loi de Poisson de paramètre 0 est la loi d'une v.a. constante et egale à 0).

La densité jointe de $(T_1,\ldots,T_n)$ est donnée par

$$f_{(T_1,\ldots,T_n)}(t_1,\ldots,t_n)=\lambda^n e^{-\lambda t_n}\mathbb{I}_{0<t_1<t_2<\cdots<t_n}$$

Calculons

$$\mathbb{P}(T_l\leqslant s\leqslant T_{l+1}\leqslant T_{l+k}\leqslant t\leqslant T_{l+k+1})=\int_A\lambda^{l+k+1}e^{-\lambda r_{l+k+1}}\mathbb{I}_{0<r_1<r_2<\cdots<r_{l+k+1}}\mathrm{d}r_1\cdots\mathrm{d}r_{l+k+1}$$

où $A=\{(r_1,\ldots,r_{l+k+1}): r_1<r_2<\cdots<r_l<s<r_{l+1}<\cdots<r_{l+k}<t<r_{l+k+1}\}$. Mais alors un calcul direct donne

$$=e^{-\lambda t}\frac{(\lambda\,s)^l}{l!}\frac{(\lambda(t-s))^k}{k!}$$

ce qu'implique

$$\mathbb{P}(N_t=l+k|N_s=l)=\mathbb{P}(T_{l+k}\leqslant t\leqslant T_{l+k+1}|T_l\leqslant s\leqslant T_{l+1})=e^{-\lambda(t-s)}\frac{(\lambda(t-s))^k}{k!}.$$

On veut maintenant montrer que la même propriété est vraie pour un nombre arbitraire d'increments. Soit $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_{n+1}$ des temps et $k_1 \leqslant \cdots \leqslant k_{n+1}$ une suite croissante d'entiers.

L'evenement $\{N_{t_1} = k_1, ..., N_{t_{n+1}} = k_{n+1}\}$ est equivalent à $\{T_{k_i} \leqslant t_i \leqslant T_{k_i+1} : i = 1, ..., n+1\}$ et

$\mathbb{P}(N_{t_{n+1}} = k_{n+1} | N_{t_n} = k_n, ..., N_{t_1} = k_1) = \mathbb{P}(T_{k_{n+1}} \leqslant t_{n+1} \leqslant T_{k_{n+1}+1} | T_{k_i} \leqslant t_i \leqslant T_{k_i+1} : i = 1, ..., n)$.

par un calcul analogue au precedent on a que

$$\mathbb{P}(T_{k_{n+1}} \leqslant t_{n+1} \leqslant T_{k_{n+1}+1} | T_{k_i} \leqslant t_i \leqslant T_{k_i+1} : i = 1, ..., n)$$

$$= \frac{\mathbb{P}(T_{k_1} \leqslant t_1 \leqslant T_{k_1+1} \leqslant T_{k_2} \leqslant t_2 \leqslant T_{k_2+1} \leqslant \cdots \leqslant T_{k_{n+1}} \leqslant t_{n+1} \leqslant T_{k_{n+1}+1})}{\mathbb{P}(T_{k_1} \leqslant t_1 \leqslant T_{k_1+1} \leqslant T_{k_2} \leqslant t_2 \leqslant T_{k_2+1} \leqslant \cdots \leqslant T_{k_n} \leqslant t_n \leqslant T_{k_n+1})}$$

$$= \frac{e^{-\lambda t_{n+1}}}{e^{-\lambda t_n}} \frac{\prod_{i=0}^{n} \frac{(\lambda(t_{i+1} - t_i))^{k_{i+1} - k_i}}{(k_{i+1} - k_i)!}}{\prod_{i=0}^{n-1} \frac{(\lambda(t_{i+1} - t_i))^{k_{i+1} - k_i}}{(k_{i+1} - k_i)!}} = e^{-\lambda(t_{n+1} - t_n)} (\lambda(t_{n+1} - t_n))^{k_{n+1} - k_n}$$

Et donc

$$\mathbb{P}(N_{t_n+h} - N_{t_n} = l | N_{t_n} = k_n, ..., N_{t_1} = k_1) = e^{-\lambda h} \frac{(\lambda h)^l}{l!} = \mathbb{P}(N_h = l)$$

qui montre que les increments du processus de Poisson sont stationnaires et ne dependent pas du passé.

Les observations precedentes nous permettent de donner une nouvelle definition du processus de Poisson, equivalente à celle de processus de comptage pour le processus de renouvellement avec intertemps exponentiels.

Définition 5. *Un processus de Poisson d'intensité $\lambda > 0$ est un processus de comptage à accroissements indépendantes et stationnaires (PAIS) tel que*

$$N_t \sim \mathcal{P}(\lambda t).$$

Pour montrer l'equivalente, il nous reste à montrer que les temps d'accroissement d'un tel processus de comptage forment un processus de renouvellement avec intertemps exponentiels. Calculons la densité de $(T_1, ..., T_n)$:

Pour $s_1 < s_2 < \cdots < s_n$ on a que

$$f_{(T_1, ..., T_n)}(s_1, ..., s_n) = \lim_{h_1, ..., h_n \to 0} \frac{\mathbb{P}(T_1 \in [s_1, s_1 + h_1[, ..., T_n \in [s_n, s_n + h_n[)}{h_1 \cdots h_n}$$

où

$$\mathbb{P}(T_1 \in [s_1, s_1 + h_1[, ..., T_n \in [s_n, s_n + h_n[)$$

$$= \mathbb{P}(N_{s_1} = 0, N_{s_1+h_1} = 1, N_{s_2} = 1, ..., N_{s_2+h_2} = 2, ..., N_{s_n+h_n} = n+1)$$

$$= e^{-\lambda s_1} \lambda h_1 e^{-\lambda h_1} \prod_{i=1}^{n-1} e^{-\lambda(s_{i+1} - s_i - h_i)} \lambda h_{i+1} e^{-\lambda h_{i+1}}.$$

Donc

$$f_{(T_1, ..., T_n)}(s_1, ..., s_n) = e^{-\lambda s_n} \lambda^n.$$

Pour les autres valeurs de $(s_1, ..., s_n)$ on a $f_{(T_1, ..., T_n)}(s_1, ..., s_n) = 0$ de sorte que

$$f_{(T_1, ..., T_n)}(s_1, ..., s_n) = e^{-\lambda s_n} \lambda^n \mathbb{I}_{0 < s_1 < \cdots < s_n}.$$

D'ici est facile voir que le vecteur $(\tau_1, \ldots, \tau_n)$ est iid de loi $\mathcal{E}(\lambda)$.

3 Propriétés du processus de Poisson

Théorème 6. *(Loi des grandes nombres) Soit $(N_t)_{t\geqslant 0}$ un processus de Poisson d'intensité λ, alors*

$$\lim_{t\to+\infty} \frac{N_t}{t} = \lambda \qquad p.s.$$

Démonstration. (1ére methode) On écrit

$$\frac{N_n}{n} = \frac{1}{n}\sum_{k=1}^{n} [N_k - N_{k-1}].$$

Par l'indépendance et la stationnarité des increments la suite $(N_k - N_{k-1})_{k\geqslant 1}$ est iid de loi $\mathcal{P}(\lambda)$. Donc par la loi forte des grandes nombres pour les suites iid on a que

$$N_n/n \to \mathbb{E}[N_1] = \lambda.$$

Or le processus $(N_t)_{t\geqslant 0}$ est croissante, donc

$$\frac{N_{\lfloor t\rfloor}}{\lfloor t\rfloor + 1} \leqslant \frac{N_{\lfloor t\rfloor}}{t} \leqslant \frac{N_t}{t} \leqslant \frac{N_{\lfloor t\rfloor+1}}{t} \leqslant \frac{N_{\lfloor t\rfloor+1}}{\lfloor t\rfloor}$$

mais maintenant $N_n/(n+1) \to \lambda$ et aussi $N_{n+1}/n \to \lambda$ et donc $N_t/t \to \lambda$. □

Démonstration. (2éme methode) Par la loi forte des grandes nombres pour les v.a. iid:

$$\frac{T_n}{n} = \frac{1}{n}\sum_{k=1}^{n} \tau_k \to \mathbb{E}[\tau_1] = \frac{1}{\lambda}$$

Mais $T_{N_t} \leqslant t \leqslant T_{N_t+1}$, donc

$$\frac{T_{N_t}}{N_t} \leqslant \frac{t}{N_t} \leqslant \frac{T_{N_t+1}}{N_t} \leqslant \frac{T_{N_t+1}}{N_t+1} \cdot \frac{N_t+1}{N_t}$$

Le processus N_t est croissante, donc $N_{+\infty} = \lim_{t\to+\infty} N_t$ existe p.s.. Maintenant il est impossible que $N_{+\infty} < +\infty$ car sinon on obtiens que $+\infty \leqslant T_{N_{+\infty}+1}$ mais $T_k < +\infty$ pour tout $k \geqslant 1$. Alors $N_{+\infty} = +\infty$ et donc on obtiens que

$$\lim_{t\to+\infty} \frac{t}{N_t} = \frac{1}{\lambda}.$$

□

Soit $(\mathcal{F}_t)_{t\geqslant 0}$ la filtration naturelle du processus $(N_t)_{t\geqslant 0}$, c-à-d $\mathcal{F}_t = \sigma(N_s\colon s \in [0, t])$. Soit $\mathcal{F}_{t+} = \cap_{s\geqslant t}\mathcal{F}_s$ et $\mathcal{F}_{t-} = \sigma(\mathcal{F}_s : s < t) = \sigma(N_s\colon 0 \leqslant s < t)$. On remarque (sans donner de preuve) que $\mathcal{F}_t = \mathcal{F}_{t+}$ et $\mathcal{F}_{t-} \neq \mathcal{F}_t$. Cela, intuitivement veut dire qu'on ne peut pas prevoir les instants de saut de N.

Théorème 7. *(Propriété de Markov) On a que, pout toute fonction $f\colon \mathbb{R} \to \mathbb{R}$ mesurable et bornée*

$$\mathbb{E}[f(N_{t+h})|\mathcal{F}_t] = (Q_h f)(N_t)$$

avec

$$Q_h f(x) = \mathbb{E}[f(x + N_h)] = \sum_{k\geqslant 0} f(x+k)\frac{(\lambda h)^k}{k!} e^{-\lambda h}$$

Démonstration. Il suffit montrer que

$$\mathbb{E}[f(N_{t+h})\, g(N_{t_1}, \ldots, N_{t_n})] = \mathbb{E}[(Q_h f)(N_t)\, g(N_{t_1}, \ldots, N_{t_n})]$$

pour tout sequence $0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant t$ car $G \hat{\in} \mathcal{F}_t$ et positive est une limite croissante de fonctions de cette forme. Mais alors

$$\mathbb{E}[f(N_{t+h})\, g(N_{t_1}, \ldots, N_{t_n})] = \mathbb{E}[\mathbb{E}[f(N_{t+h} - N_t + N_t) | N_t, N_{t_1}, \ldots, N_{t_n}]\, g(N_{t_1}, \ldots, N_{t_n})]$$

et

$$\mathbb{E}[f(N_{t+h} - N_t + N_t) | N_t, N_{t_1}, \ldots, N_{t_n}] = \varphi(N_t)$$

où $\varphi(x) = \mathbb{E}[f(N_{t+h} - N_t + x)]$ car la v.a. $N_{t+h} - N_t$ est indépendante de $(N_{t_1}, \ldots, N_{t_n}, N_t)$. Par stationnarité $\varphi(x) = \mathbb{E}[f(N_h + x)] = Q_h f(x)$. □

Exercice 1. Montrer que $Q_h Q_r f = Q_{r+h} f$ pour tout $h, r \geqslant 0$.

Théorème 8. *(Somme de PP) La somme de deux processus de Poisson $N^{(1)}, N^{(2)}$ indépendants d'intensités respectives λ_1 et λ_2 est un processus de Poisson d'intensité $\lambda_1 + \lambda_2$.*

Démonstration. Le processus $N_t = N_t^{(1)} + N_t^{(2)}$ est à increments indépendants et stationnaires et la loi de N_t est $\mathcal{P}((\lambda_1 + \lambda_2)t)$ étant la somme de deux lois de Poisson de paramètres $\lambda_1 t$ et $\lambda_2 t$. □

Soit $(X_1, \ldots, X_n)$ un n-uplet de v.a. réelles. On note $(X_{(1)}, \ldots, X_{(n)})$ les statistiques d'ordre des $(X_1, \ldots, X_n)$, c-à-d la suite des valeurs de $X_1, \ldots, X_n$ ordonnées de façon croissante: $X_{(1)} \leqslant X_{(2)} \leqslant X_{(3)} \leqslant \cdots \leqslant X_{(n)}$ et $\{X_{(i)}\}_{i=1,\ldots,n} = \{X_i\}_{i=1,\ldots,n}$.

Un vecteur aléatoire $(X_1, \ldots, X_n)$ est *échangeable* ssi $(X_1, \ldots, X_n) \sim (X_{\sigma(1)}, \ldots, X_{\sigma(n)})$ pour tout $\sigma \in S_n$ (permutations de n objects). Un vecteur iid est toujours échangeable. La densité d'un vecteur échangeable est une fonction symétrique de ses arguments.

Théorème 9. *Soit $(X_1, \ldots, X_n)$ un vecteur aléatoire échangeable de densité $f_{(X_1, \ldots, X_n)}(x_1, \ldots, x_n)$ alors le vecteur des statistiques d'ordre $(X_{(1)}, \ldots, X_{(n)})$ a densité*

$$f_{(X_{(1)}, \ldots, X_{(n)})}(x_1, \ldots, x_n) = n!\, f_{(X_1, \ldots, X_n)}(x_1, \ldots, x_n)\, \mathbb{I}_{x_1 < x_2 < \cdots < x_n}.$$

Démonstration. On remarque d'abord que pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n$ tel que $x_i \neq x_j$ pour tout $1 \leqslant i < j < n$ il existe une seule permutation $\sigma \in S_n$ telle que $x_{\sigma(1)} < x_{\sigma(2)} < \cdots < x_{\sigma(n)}$ et en particulier que

$$\mathbb{I}_{\{x_i \neq x_j, 1 \leqslant i < j \leqslant n\}} = \sum_{\sigma \in S_n} \mathbb{I}_{x_{\sigma(1)} < x_{\sigma(2)} < \cdots < x_{\sigma(n)}}.$$

Maintenant

$$\mathbb{E}[\varphi(X_{(1)}, \ldots, X_{(n)})] = \int_{\mathbb{R}^n} \varphi(x_{(1)}, \ldots, x_{(n)}) f_{(X_1, \ldots, X_n)}(x_1, \ldots, x_n) \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

$$= \int_{\mathbb{R}^n} \varphi(x_{(1)}, \ldots, x_{(n)}) f_{(X_1, \ldots, X_n)}(x_1, \ldots, x_n) \mathbb{I}_{\{x_i \neq x_j, 1 \leqslant i < j \leqslant n\}} \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

$$= \sum_{\sigma \in S_n} \int_{\mathbb{R}^n} \varphi(x_{(1)}, \ldots, x_{(n)}) f_{(X_1, \ldots, X_n)}(x_1, \ldots, x_n) \mathbb{I}_{x_{\sigma(1)} < x_{\sigma(2)} < \cdots < x_{\sigma(n)}} \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

mais sur l'ensemble $\{x_{\sigma(1)} < \cdots < x_{\sigma(n)}\}$ on a que exactement $x_{(i)} = x_{\sigma(i)}$ pour tout $i = 1, \ldots, n$:

$$= \sum_{\sigma \in S_n} \int_{\mathbb{R}^n} \varphi(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) f_{(X_1, \ldots, X_n)}(x_1, \ldots, x_n) \mathbb{I}_{x_{\sigma(1)} < x_{\sigma(2)} < \cdots < x_{\sigma(n)}} \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

changement de variables $y_i = x_{\sigma(i)}$ $\left(x_i = y_{\sigma^{-1}(i)}\right)$

$$= \sum_{\sigma \in S_n} \int_{\mathbb{R}^n} \varphi(y_1, \ldots, y_n) f_{(X_1, \ldots, X_n)}\left(y_{\sigma^{-1}(1)}, \ldots, y_{\sigma^{-1}(n)}\right) \mathbb{I}_{y_1 < y_2 < \cdots < y_n} \mathrm{d}y_1 \cdots \mathrm{d}y_n$$

par echangeabilité

$$= \sum_{\sigma \in S_n} \int_{\mathbb{R}^n} \varphi(y_1, \ldots, y_n) f_{(X_1, \ldots, X_n)}(y_1, \ldots, y_n) \mathbb{I}_{y_1 < y_2 < \cdots < y_n} \mathrm{d}y_1 \cdots \mathrm{d}y_n$$

comme rien plus depends de la permutation σ on peut calculer la somme et obtenir

$$= n! \int_{\mathbb{R}^n} \varphi(y_1, \ldots, y_n) f_{(X_1, \ldots, X_n)}(y_1, \ldots, y_n) \mathbb{I}_{y_1 < y_2 < \cdots < y_n} \mathrm{d}y_1 \cdots \mathrm{d}y_n$$

ce qu'implique que la densité des statistiques d'ordre est

$$f_{(X_{(1)}, \ldots, X_{(n)})}(x_1, \ldots, x_n) = n! \, f_{(X_1, \ldots, X_n)}(x_1, \ldots, x_n) \mathbb{I}_{x_1 < \cdots < x_n}.$$

□

Théorème 10. *(Statistiques d'ordre) Conditionnellement à l'evenement $\{N_t = n\}$ le vecteur aléatoire $(T_1, \ldots, T_n)$ a la même loi des statistiques d'ordre $(U_{(1)}, \ldots, U_{(n)})$ d'un n-uplet $(U_1, \ldots, U_n)$ de v.a. iid uniformes sur $[0, t]$.*

Démonstration. Calculons la loi de $(T_1, \ldots, T_n)$ conditionnellement à $\{N_t = n\}$: soit $(s_1, \ldots, s_n) \in \mathbb{R}_+^n$ et $h_1, \ldots, h_n$ des quantités positives et petites: la probabilité

$$F(s_1, \ldots, s_n) = \mathbb{P}(T_1 \leqslant s_1, \ldots, T_n \leqslant s_n | N_t = n) = \frac{\mathbb{P}(T_1 \leqslant s_1, \ldots, T_n \leqslant s_n, N_t = n)}{\mathbb{P}(N_t = n)}$$

$$= \frac{\mathbb{P}(T_1 \leqslant s_1, \ldots, T_n \leqslant s_n, T_n \leqslant t < T_{n+1})}{(\lambda t)^n / n! e^{-\lambda t}}$$

est donnée par

$$F(s_1, \ldots, s_n) = \frac{n! e^{\lambda t}}{(\lambda t)^n} \int_0^{s_1} \mathrm{d}t_1 \cdots \int_0^{s_n \wedge t} \mathrm{d}t_n \int_t^{+\infty} \mathrm{d}t_{n+1} f_{(T_1, \ldots, T_{n+1})}(t_1, \ldots, t_{n+1})$$

$$= \frac{n! e^{\lambda t}}{(\lambda t)^n} \int_0^{s_1} \mathrm{d}t_1 \cdots \int_0^{s_n \wedge t} \mathrm{d}t_n \int_t^{+\infty} \mathrm{d}t_{n+1} \lambda^{n+1} e^{-\lambda t_{n+1}} \mathbb{I}_{t_1 < t_2 < \cdots < t_n < t_{n+1}}$$

$$= \frac{1}{t^n} \int_0^{s_1} \mathrm{d}t_1 \cdots \int_0^{s_n \wedge t} \mathrm{d}t_n \mathbb{I}_{t_1 < t_2 < \cdots < t_n}$$

On obtiens la densité conditionnelle $f_{(T_1, \ldots, T_n)|N_t = n}(t_1, \ldots, t_n)$ de $(T_1, \ldots, T_n)$ en prenant les dérivées par rapport à $s_1, \ldots, s_n$ de l'expression precedente:

$$f_{(T_1, \ldots, T_n)|N_t = n}(t_1, \ldots, t_n) = \frac{\partial^n}{\partial t_1 \cdots \partial t_n} F(t_1, \ldots, t_n) = \frac{n!}{t^n} \mathbb{I}_{0 < t_1 < \cdots < t_n < t}$$

qui est précisément la densité des statistiques d'ordre d'un n-plet des v.a. uniformes sur $[0, t]$:

$$f_{(T_1, \ldots, T_n)|N_t = n}(t_1, \ldots, t_n) = n! \mathbb{I}_{0 < t_1 < \cdots < t_n < t} \, f_{(U_1, \ldots, U_n)}(t_1, \ldots, t_n)$$

avec $f_{(U_1, \ldots, U_n)}(t_1, \ldots, t_n) = \prod_{i=1}^n f_{\mathcal{U}([0,t])}(t_i)$. □

En particulier on a que

$$\mathbb{E}[\varphi(T_{(1)}, \ldots, T_{(n)}) | N_t = n] = \mathbb{E}[\varphi(U_{(1)}, \ldots, U_{(n)})]$$

avec $U_i \sim \mathcal{U}([0,t])$ et iid.

Exemple 11. Les employés d'un entreprise arrivent au travail selon un processus de Poisson d'intensité $\lambda > 0$. Pour tout $n \geqslant 1$ soit T_n le temps d'arrivée du n-éme employé. Le temps total de travail effectuée par les employés jusqu'à l'instant t est donné par

$$T_t = \sum_{k \geqslant 1} (t - T_k)_+ .$$

On s'interesse à la moyenne de cette variable aléatoire:

$$\mathbb{E}[T_t] = \mathbb{E}\left[\sum_{k \geqslant 1} (t - T_k)_+\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{k \geqslant 1} (t - T_k)_+ \middle| N_t\right]\right].$$

Or

$$\mathbb{E}\left[\sum_{k \geqslant 1} (t - T_k)_+ \middle| N_t = n\right] = \mathbb{E}\left[\sum_{k=1}^{n} (t - T_k)_+ \middle| N_t = n\right] = \mathbb{E}\left[\sum_{k=1}^{n} (t - T_{(k)})_+ \middle| N_t = n\right]$$

$$= \mathbb{E}\left[\sum_{k=1}^{n} (t - U_{(k)})_+\right] = \mathbb{E}\left[\sum_{k=1}^{n} (t - U_k)_+\right] = \sum_{k=1}^{n} \mathbb{E}[(t - U_k)_+] = n\mathbb{E}[(t - U_1)_+]$$

et

$$\mathbb{E}[(t - U_1)_+] = \int_0^t (t - x)\frac{\mathrm{d}x}{t} = \int_0^t x\frac{\mathrm{d}x}{t} = \frac{t}{2}.$$

Finalement on obtient $\mathbb{E}[T_t] = t/2\,\mathbb{E}[N_t] = \lambda t^2/2$.

4 Mélange de Processus de Poisson (Processus de Cox)

Du point de vue des applications (à l'actuariat, notamment), la modelisation des temps d'arrivée par un processus de Poisson peut se reveler simpliste. Par, exemple si on s'interesse aux temps d'arrivée d'accidents de voiture, la frequence moyenne d'accidents peut differer fortement d'une persone à l'autre, dependant de l'age d'experience, de la façon de conduire, du type de voiture. On essaye de prendre en compte cette variabilité en introduisant une intensité aléatoire dans le modele de Poisson, c-à-d on fait dependre l'intensité λ elle même de l'experience aléatoire $\omega \in \Omega$. Il est aussi un cas particulier d'une classe de procesus appelés processus de Cox [Cox, D. R. (1955). "Some Statistical Methods Connected with Series of Events". Journal of the Royal Statistical Society 17 (2): 129-164.]

Définition 12. *Soit* $N \sim \mathrm{PP}(\lambda)$ *et* Θ *une v.a. strictement positive independante de N. Le processus* $\tilde{N}_t = N_{\Theta t}$ *est appelé un mélange de* $\mathrm{PP}(\lambda)$ *de loi mélangeante* Θ.

Un point de vue intuitif est que $\tilde{N}$ est un procesus de Poisson d'intensité aléatoire donnée par $\lambda\Theta > 0$.

Quelques proprietes de $\tilde{N}$:

1. Moyenne

$$\mathbb{E}\big[\tilde{N}_t|\Theta\big] = \mathbb{E}[N_{\Theta t}|\Theta] = \lambda\,\Theta\,t, \qquad \mathbb{E}\big[\tilde{N}_t\big] = \mathbb{E}[\mathbb{E}[N_{\Theta t}|\Theta]] = \lambda\,t\,\mathbb{E}[\Theta]$$

2. Variance

$$\mathrm{Var}\big(\tilde{N}_t|\Theta\big) = \lambda\,\Theta\,t$$

$$\mathrm{Var}\big(\tilde{N}_t\big) = \mathrm{Var}\big(\mathbb{E}\big[\tilde{N}_t|\Theta\big]\big) + \mathbb{E}\big[\mathrm{Var}\big(\tilde{N}_t|\Theta\big)\big] = \lambda^2\,t^2\,\mathrm{Var}(\Theta) + \lambda\,t\,\mathbb{E}[\Theta]$$

3. Les increments ne sont pas independants:

$$\mathrm{Cov}\big(\tilde{N}_t - \tilde{N}_s, \tilde{N}_s\big) \neq 0$$

(exercice)

4. Les increments sont stationnaires

$$\mathbb{E}\big[\, f\big(\tilde{N}_t - \tilde{N}_s\big)\big] = \mathbb{E}[\mathbb{E}[f(N_{\Theta t} - N_{\Theta s})|\Theta]] = \mathbb{E}[\mathbb{E}[f(N_{\Theta t - \Theta s})|\Theta]] = \mathbb{E}\big[\, f\big(\tilde{N}_{t-s}\big)\big]\,.$$

Deux autres proprietés importantes sont liées aux statistique d'ordre des temps de saut et à la preservation du caractere Markovien du processus.

Théorème 13. *Conditionnelement à $\{\tilde{N}_t = n\}$ la loi des temps de saut $(T_1, \ldots, T_n)$ est independante de Θ et donnée par les statistiques d'ordre d'un n-uplet de v.a. uniformes sur $[0, t]$.*

Démonstration. Conditionellement à Θ le processus est de Poisson d'intensité $\lambda\Theta$ donc par les resultats sur la loi des temps de saut d'un processus de Poisson on obtient que

$$\mathbb{E}[f(T_1, \ldots, T_n)\mathbb{I}_{T_n \leqslant t < T_{n+1}}|\Theta] =$$

$$= (\lambda\Theta)^{n+1} \int_{0 < t_1 < \cdots < t_{n+1}} e^{-\lambda\Theta t_{n+1}} f(t_1, \ldots, t_n)\mathbb{I}_{t_n \leqslant t < t_{n+1}} \mathrm{d}t_1 \cdots \mathrm{d}t_{n+1}$$

$$= (\lambda\Theta)^n e^{-\lambda\Theta t} \int_{0 < t_1 < \cdots < t_n \leqslant t} f(t_1, \ldots, t_n) \mathrm{d}t_1 \cdots \mathrm{d}t_{n+1}$$

Donc

$$\mathbb{E}\big[\, f(T_1, \ldots, T_n)\, g(\Theta)\big|\tilde{N}_t = n\big] = \frac{\mathbb{E}\big[\, f(T_1, \ldots, T_n) g(\Theta)\mathbb{I}_{\tilde{N}_t = n}\big]}{\mathbb{P}\big(\tilde{N}_t = n\big)}$$

$$= \frac{\mathbb{E}[f(T_1, \ldots, T_n) g(\Theta)\mathbb{I}_{T_n \leqslant t < T_{n+1}}]}{\mathbb{E}[\mathbb{I}_{T_n \leqslant t < T_{n+1}}]}$$

$$= \frac{\mathbb{E}[\mathbb{E}[f(T_1, \ldots, T_n)\mathbb{I}_{T_n \leqslant t < T_{n+1}}|\Theta] g(\Theta)]}{\mathbb{E}[\mathbb{I}_{T_n \leqslant t < T_{n+1}}]}$$

$$= \frac{\mathbb{E}[(\lambda\Theta)^n e^{-\lambda\Theta t} g(\Theta)]}{\mathbb{E}[(\lambda\Theta)^n e^{-\lambda\Theta t}]} \frac{n!}{t^n} \int_{0 < t_1 < \cdots < t_n \leqslant t} f(t_1, \ldots, t_n) \mathrm{d}t_1 \cdots \mathrm{d}t_{n+1}$$

$$= \frac{\mathbb{E}[(\lambda\Theta)^n e^{-\lambda\Theta t} g(\Theta)]}{\mathbb{E}[(\lambda\Theta)^n e^{-\lambda\Theta t}]} \frac{n!}{t^n} \int_{0 < t_1 < \cdots < t_n \leqslant t} f(t_1, \ldots, t_n) \mathrm{d}t_1 \cdots \mathrm{d}t_{n+1}$$

$$== \frac{\mathbb{E}[(\lambda\Theta)^n e^{-\lambda\Theta t} g(\Theta)]}{\mathbb{E}[(\lambda\Theta)^n e^{-\lambda\Theta t}]} \mathbb{E}[f(U_{(1)}, \ldots, U_{(n)})]$$

Si on pose $f = 1$ on a aussi que

$$\mathbb{E}\big[\, g(\Theta)\big|\tilde{N}_t = n\big] = \frac{\mathbb{E}[(\lambda\Theta)^n e^{-\lambda\Theta t} g(\Theta)]}{\mathbb{E}[(\lambda\Theta)^n e^{-\lambda\Theta t}]}$$

et donc

$$\mathbb{E}\big[\, f(T_1, \ldots, T_n)\, g(\Theta)\big|\tilde{N}_t = n\big] = \mathbb{E}\big[\, g(\Theta)\big|\tilde{N}_t = n\big]\mathbb{E}[f(U_{(1)}, \ldots, U_{(n)})]$$

ce qui montre l'independance conditionelle de Θ et $(T_1, \ldots, T_n)$ sachant $\{\tilde{N}_t = n\}$ et qui donne que conditionellement à cet evenement la loi de $(T_1, \ldots, T_n)$ est celle de $(U_{(1)}, \ldots, U_{(n)})$ avec $(U_1, \ldots, U_n)$ un n-uple iid uniforme sur $[0, t]$. □

Corollaire 14. *Soit* $0 \leqslant t_1 \leqslant \cdots \leqslant t_n$ *des temps alors la loi de* $(\tilde{N}_{t_1}, \ldots, \tilde{N}_{t_n})$ *conditionellement à* $\tilde{N}_{t_n} = m$ *ne depends pas de la loi de* Θ *ou de* λ *et elle est egale à la même loi pour n'importe quel autre processus de Poisson d'intensité arbitraire* $\mu > 0$.

Démonstration. Calculons $\mathbb{P}(\tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n)$: par le resultat precedent la loi de $(T_1, \ldots, T_{k_n})$ conditionellement à $\{\tilde{N}_{t_n} = k_n\}$ à la même loi des premiers k_n temps de saut d'un processus de Poisson d'intensité $\mu > 0$ (en effet la loi ne depends pas de l'intensité), donc:

$$\mathbb{P}(\tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n | \tilde{N}_{t_n} = k_n) = \mathbb{P}(T_{k_1} \leqslant t_1 < T_{k_1+1}, \ldots, T_{k_{n-1}} \leqslant t_{n-1} < T_{k_{n-1}+1} | \tilde{N}_{t_n} = k_n)$$

$$= \mathbb{P}(N_{t_1} = k_1, \ldots, N_{t_n} = k_n | N_{t_n} = k_n)$$

□

Cette derniere proprieté est la clé pour demontrer la proprieté de Markov de $(\tilde{N}_t)_{t \geqslant 0}$.

Théorème 15. $(\tilde{N}_t)_{t \geqslant 0}$ *est un processus de Markov par rapport à sa filtration naturelle.*

Démonstration. Donc

$$\mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1} | \tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n) = \frac{\mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1}, \tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n)}{\mathbb{P}(\tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n)}$$

$$= \frac{\mathbb{P}(\tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n | \tilde{N}_{t_{n+1}} = k_{n+1})}{\mathbb{P}(\tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_{n-1}} = k_{n-1} | \tilde{N}_{t_n} = k_n)} \frac{\mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1})}{\mathbb{P}(\tilde{N}_{t_n} = k_n)}$$

$$= \frac{\mathbb{P}(N_{t_1} = k_1, \ldots, N_{t_n} = k_n | N_{t_{n+1}} = k_{n+1})}{\mathbb{P}(N_{t_1} = k_1, \ldots, N_{t_{n-1}} = k_{n-1} | N_{t_n} = k_n)} \frac{\mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1})}{\mathbb{P}(\tilde{N}_{t_n} = k_n)}$$

$$= \frac{\mathbb{P}(N_{t_1} = k_1, \ldots, N_{t_n} = k_n, N_{t_{n+1}} = k_{n+1})}{\mathbb{P}(N_{t_1} = k_1, \ldots, N_{t_{n-1}} = k_{n-1}, N_{t_n} = k_n)} \frac{\mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1})}{\mathbb{P}(\tilde{N}_{t_n} = k_n)} \frac{\mathbb{P}(N_{t_n} = k_n)}{\mathbb{P}(N_{t_{n+1}} = k_{n+1})}$$

$$= \mathbb{P}(N_{t_{n+1}} = k_{n+1} | N_{t_1} = k_1, \ldots, N_{t_n} = k_n) \frac{\mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1})}{\mathbb{P}(\tilde{N}_{t_n} = k_n)} \frac{\mathbb{P}(N_{t_n} = k_n)}{\mathbb{P}(N_{t_{n+1}} = k_{n+1})}$$

$$\mathbb{P}(N_{t_{n+1}} = k_{n+1} | N_{t_n} = k_n) \frac{\mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1})}{\mathbb{P}(\tilde{N}_{t_n} = k_n)} \frac{\mathbb{P}(N_{t_n} = k_n)}{\mathbb{P}(N_{t_{n+1}} = k_{n+1})}$$

qui montre que $(k_1, \ldots, k_{n+1}) \mapsto \mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1} | \tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n)$ est une fonction des seules variables k_n, k_{n+1} ce qui implique facilement que

$$\mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1} | \tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n) = \mathbb{P}(\tilde{N}_{t_{n+1}} = k_{n+1} | \tilde{N}_{t_n} = k_n)$$

cela implique que pour tout $t_1 < \cdots < t_n = s < t$ on a que

$$\mathbb{E}[f(\tilde{N}_t) g(\tilde{N}_{t_1}, \ldots, \tilde{N}_{t_n})] = \mathbb{E}[\mathbb{E}[f(\tilde{N}_t) | \tilde{N}_{t_1}, \ldots, \tilde{N}_{t_n}] g(\tilde{N}_{t_1}, \ldots, \tilde{N}_{t_n})]$$

$$= \mathbb{E}[\mathbb{E}[f(\tilde{N}_t) | \tilde{N}_{t_n}] g(\tilde{N}_{t_1}, \ldots, \tilde{N}_{t_n})] = \mathbb{E}[\mathbb{E}[f(\tilde{N}_t) | \tilde{N}_s] g(\tilde{N}_{t_1}, \ldots, \tilde{N}_{t_n})]$$

mais comme toute fonction positive bornée et mesurable par rapport à $\mathcal{F}_s$ peut s'approcher de façon monotone par des fonction de la forme $g(\tilde{N}_{t_1}, \ldots, \tilde{N}_{t_n})$ on obtiens que

$$\mathbb{E}[f(\tilde{N}_t) G] = \mathbb{E}[\mathbb{E}[f(\tilde{N}_t) | \tilde{N}_s] G]$$

pour tout G bornée et $\mathcal{F}_s$-mesurable et donc que $\mathbb{E}\big[f\big(\tilde{N}_t\big)|\mathcal{F}_s\big] = \mathbb{E}\big[f\big(\tilde{N}_t\big)|\tilde{N}_s\big]$, c-à-d la proprieté de Markov en temps continu. □

Exemple 16. Un assureur pense que chacun de ses assurée va demander des remboursements selons un processus de Poisson d'intensité $\lambda \in [0, 1]$ (mesuré en remboursements/année). Il pense aussi que chaque assurée soit caracterisé par une propre intensité λ et que chaque nouveau assuré puisse être associé à une intensité uniformement distribuée sur $[0, 1]$. Si l'assuré à demande n remboursements dans ses premieres t années quel est la loi conditionelle du temps restant jusqu'à la prochaine demande de remboursement de la part de cet assuré?

Solution. Le processus des remboursement est modelisé par un mélange de processus de Poisson d'intensité 1 et loi melangeanté $\Theta \sim \mathcal{U}([0, 1])$. Conditionellement à $\tilde{N}_t = n$ le prochain evenement a lieu à l'instant T_{n+1}, donc la probabilité cherchée est

$$\mathbb{P}\big(T_{n+1} > x|\tilde{N}_t = n\big) = \frac{\mathbb{E}\Big[\mathbb{I}_{T_{n+1}>x}\mathbb{I}_{\tilde{N}_t=n}\Big]}{\mathbb{E}\Big[\mathbb{I}_{\tilde{N}_t=n}\Big]} = \frac{\mathbb{E}\big[\mathbb{E}\big[\mathbb{I}_{T_{n+1}>x}\mathbb{I}_{\tilde{N}_t=n}|\Theta\big]\big]}{\mathbb{E}\Big[\mathbb{E}\Big[\mathbb{I}_{\tilde{N}_t=n}|\Theta\Big]\Big]}$$

$$= \frac{\mathbb{E}\big[\mathbb{E}\big[\mathbb{I}_{T_{n+1}>x}\mathbb{I}_{\tilde{N}_t=n}|\Theta\big]\big]}{\mathbb{E}\Big[\mathbb{E}\Big[\mathbb{I}_{\tilde{N}_t=n}|\Theta\Big]\Big]}$$

maintenant

$$\mathbb{P}(T_{n+1} > x, N_{\theta t} = n) = \mathbb{P}(T_{n+1} > x|N_{\theta t} = n)\mathbb{P}(N_{\theta t} = n) = \int_x^{+\infty} \theta e^{-\theta y}\mathrm{d}y\,\frac{(\theta t)^n}{n!}e^{-\theta t}$$

$$= e^{-\theta(x+t)}\frac{(\theta t)^n}{n!}e^{-\theta t}$$

et donc

$$\mathbb{P}\big(T_{n+1} > x|\tilde{N}_t = n\big) = \frac{\mathbb{E}\Big[e^{-\Theta\,(x+t)}\frac{(\Theta t)^n}{n!}\Big]}{\mathbb{E}\Big[\frac{(\Theta t)^n}{n!}e^{-\Theta t}\Big]} = \frac{\mathbb{E}\big[e^{-\Theta\,(x+t)}\,(\Theta t)^n\big]}{\mathbb{E}[e^{-\Theta\,t}\,(\Theta t)^n]} = \frac{\int_0^1 e^{-\theta(x+t)}\theta^n\mathrm{d}\theta}{\int_0^1 e^{-\theta t}\theta^n\mathrm{d}\theta}.$$