LES PROCESSUS DE POISSON

1°/ Loi exponentielle : absence de mémoire.

Soit T une variable aléatoire positive ou nulle, et $F(t)$ sa fonction de répartition. Il sera commode d'interpréter T comme la durée d'un certain phénomène : par exemple, durée de vie d'un tube au néon, d'une particule instable etc.. Sachant que le phénomène a déjà duré un temps t_0, on s'interesse à sa durée résiduelle $T - t_0$. La loi de cette durée résiduelle est la loi de $T-t_0$ conditionnelle en $T \geqslant t_0$. Notons la $F_{t_0}(t)$. D'après la définition des probabilités conditionnelles, on a :

$$1 - F_{t_0}(t) = \frac{P(T-t_0 \geqslant t \text{ et } T \geqslant t_0)}{P(T \geqslant t_0)} = \frac{P(T \geqslant t+t_0)}{P(T \geqslant t_0)}$$

soit :

$$(1) \qquad 1 - F_{t_0}(t) = \frac{1-F(t+t_0)}{1-F(t_0)}$$

Pour que la durée de vie résiduelle obéisse à la même loi que la durée de vie totale, soit $F_{t_0}(t) = F(t)$, il faut et il suffit que $F(t)$ soit une loi exponentielle, soit $1 - F(t) = e^{-at}$ $(a > 0)$

En effet, si $1 - F(t) = e^{-at}$, on a

$$1 - F_{t_0}(t) = \frac{e^{-a(t_0+t)}}{e^{-at_0}} = e^{-at} = 1-F(t)$$

Inversement, si $F_{t_0}(t) = F(t)$, la fonction $H(t) = 1 - F(t)$ vérifie :

$$H(t + t_0) = H(t)\,H(t_0)$$

On sait que les seules fonctions non croissantes vérifiant cette relation fonctionnelle sont les fonctions $H(t) = e^{-at}$ $(a \geqslant 0)$. Ici, $a = 0$ est exclu, puisque $1 - H(t)$ doit être une fonction de répartition.

Cette propriété d'absence de mémoire caractérise donc la loi exponentielle. (exemples: en physique, la loi de la radioactivité, etc...)

Sous forme différentielle, cette propriété se présente comme suit : soit $\delta t > 0$ très petit. La probabilité pour que le phénomène s'arrête entre t et $t + \delta t$ sachant qu'il durait encore au temps t est

$$1 - e^{-a\delta t} = a\delta t + O(\delta t)$$

$O(\delta t)$ désignant un infiniment petit d'ordre supérieur à 1 en δt.

Inversement, la relation :

$$(2) \qquad 1 - F_{t_0}(\delta t) = a\,\delta t + O(\delta t)$$

s'écrit :

$$H(t_0 + \delta t) - H(t_0) = - a\, H(t_0)\,\delta t + O(\delta t)$$

avec $H(t) = 1 - F(t)$, et montre que $\log H(t)$ admet une dérivée à droite constante égale à $-a$, d'où résulte encore

$$H(t) = e^{-at}$$

Ainsi, la loi exponentielle est encore caractérisée par le fait que la probabilité d'arrêt admet une densité conditionnelle constante a.

Remarque. Pour toute variable $X \geqslant 0$ dont la loi vérifie $F(t) < 1$ pour tout $t < \infty$, on peut trouver une représentation de la forme :

$$1 - F(t) = e^{-\int_0^t \lambda(dt)}$$

avec une mesure positive $\lambda(dt)$ que l'on peut encore interpréter comme une densité conditionnelle (non constante) d'arrêt. Il suffit de prendre :

$$\lambda(dt) = \frac{F(dt)}{1 - F(t)}$$

La relation $\lambda(dt) = a\,dt$ caractérise alors la loi exponentielle.

2º/ Processus de Poisson : homogénéité et absence de mémoire.

Nous considérons maintenant un phénomène dont la manifestation est instantanée et se reproduit à des instants aléatoires successifs T_1, T_2 ... (exemples : enregistrement d'une particule dans un compteur, arrivée d'un client dans une file d'attente etc..). Nous caractérisons ce phénomène par les intervalles de temps :

$$S_1 = T_1$$
$$S_2 = T_2 - T_1$$
$$\dots\dots\dots\dots\dots$$
$$S_n = T_n - T_{n-1}.$$
$$\dots\dots\dots\dots\dots$$

séparant deux manifestations successives. Nous dirons qu'il s'agit d'un processus de Poisson si ces intervalles S_n possèdent les deux propriétés suivantes :

1/ Les S_n sont des variables mutuellement indépendantes
2/ Les S_n obéissent à une même loi exponentielle de densité $a\,e^{-at}$

Ce processus possède alors la propriété fondamentale suivante :

Soit $t_0 > 0$ un instant quelconque, T'_1, T'_2 ... les manifestations successives du phénomène postérieures à t. Les $T'_1 - t_0$, $T'_2 - t_0$ ... constituent un nouveau processus de Poisson, admettant même loi que le processus initial, et indépendant des évènements antérieurs à t_0. Autrement dit, tout se passe comme si le phénomène repartait à zéro à chaque instant t, en oubliant son passé.

Celà résulte aussitot de l'absence de mémoire de la loi exponentielle. $T'_1 - t_0$, en effet, obéit encore à la loi de densité $a\,e^{-at}$ quels qu'aient pu être les évènements antérieurs à t_0, et les $T'_2 - T'_1$, $T'_3 - T'_2$ ... sont par hypothèse indépendants et soumis à cette même loi.

Nous résumerons cette propriété capitale en disant que le processus de Poisson est homogène (dans le temps) et sans mémoire.

3°/ Loi des temps d'attente.

Par hypothèse, $T_1 = S_1$ obéit à la loi de densité $a\,e^{-at}$ admettant la transformée de Laplace $a/(a+\lambda)$. T_n, instant de la n^{me} manifestation, ou aussi bien temps d'attente (compté à partir de $t = 0$) de cette n^{me} manifestation, se présente sous la forme :

$$T_n = S_1 + S_2 + \dots + S_n$$

de la somme de n variables indépendantes de même loi exponentielle. Par suite, T_n admet la transformée de Laplace :

$$\Phi_n(\lambda) = \left(\frac{a}{a+\lambda}\right)^n$$

et sa loi $F_n(t)$ est la loi gamma de paramètre $\alpha = n$, dont la densité est :

(3)
$$f_n(t) = \frac{a^n t^{n-1}}{(n-1)!} e^{-at}$$

4º/ Loi du nombre $N(t)$ des manifestations entre les instants 0 et t.

Soit $N(t)$ le nombre (aléatoire) des manifestations du phénomène entre les instants 0 et t. L'équivalence :

$$N(t) < n \iff T_n > t$$

donne :

$$P(N(t) < n) = \int_t^{\infty} f_n(t)\,dt = \int_t^{\infty} \frac{a^n t^{n-1}}{(n-1)!} e^{-at}\,dt$$

Posons :

$$h_n(t) = P[N(t) = n] = P[N(t) < n+1] - P[N(t) < n]$$

On a :

$$h_n(t) = \int_t^{\infty} \left[\frac{a^{n+1} t^n}{n!} - \frac{a^n t^{n-1}}{(n-1)!} \right] e^{-at} = -\int_t^{\infty} \frac{d}{dt} \left(\frac{a^n t^n e^{-at}}{n!} \right) dt$$

D'où la loi de $N(t)$, qui est la loi de Poisson de paramètre at :

$$(4) \qquad h_n(t) = \frac{(at)^n}{n!} e^{-at}$$

et sa fonction génératrice :

$$(5) \qquad G(s,t) = \sum_{n=0}^{\infty} s^n h_n(t) = e^{-at(1-s)}$$

Soient maintenant $I_1, I_2, \ldots I_k$ des intervalles de temps disjoints ou possèdant au plus un point commun. Si I_j est l'intervalle (t_j, t'_j), on posera :

$$N(I_j) = N(t'_j) - N(t_j)$$

(nombre des manifestations dans l'intervalle I_j). Les $N(I_j)$ sont alors des variables indépendantes, et, si L_j est la longueur de l'intervalle I_j $N(I_j)$ obéit à la loi de Poisson de paramètre aL_j .

Celà résulte immédiatement du fait que le processus est homogène et sans mémoire.

5º/ Caractérisation des processus de Poisson.

Le processus de Poisson possède les trois propriétés suivantes :

a/ Homogénéité : la variable $N(t_o, t_o + t) = N(t_o + t) - N(t)$ obéit à la même loi que $N(t)$ quel que soit $t_o > 0$.

b/ Absence de mémoire : si les I_j sont des intervalles disjoints, ou présentant au plus un point commun, les $N(I_j)$ sont mutuellement indépendantes.

c/ Pour δt petit, on a

$$\begin{cases} p_0(\delta t) = P[N(\delta t) = 0] = 1 - a\,\delta t + O(\delta t) \\ p_1(\delta t) = P[N(\delta t) = 1] = a\,\delta t + O(\delta t) \end{cases}$$

les $O(\delta t)$ designent des infiniment petits d'ordre supérieurs à 1, et a une constante positive.

Cette propriété c/ découle immédiatement de l'expression (4) des $p_n(t)$. Elle entraine la conséquence suivante :

$$c'/ \quad P[N(\delta t) > 1] = O(\delta t)$$

autrement dit, la probabilité pour qu'il y ait plus d'une manifestation entre 0 et δt est un infiniment petit d'ordre supérieur à 1.

Inversement, d'ailleurs, la propriété c'/, jointe à a/ et b/ entraine c/, de sorte que le système a/, b/ et c/ est équivalent au système a'/, b/ et c'/.

En effet, a/ et b/ entrainent :

$$p_0(t + t') = p_0(t)\, p_0(t')$$

d'où $p_0(t) = e^{-at}$ avec une constante $a > 0$, et par suite :

$$p_0(\delta t) = 1 - a\,\delta t + O(\delta t)$$

Dans ces conditions, c'/ entraine $p_1(\delta t) = a\,\delta t + O(\delta t)$ et c/ est vérifié.

Réciproquement, ces trois propriétés a/, b/ et c/ (ou a/, b/ et c'/) caractérisent les processus de Poisson.

Soit, en effet, un phénomène se manifestant à des instants successifs T_1, T_2... tels que les nombres $N(I_j)$ de manifestations observées dans des intervalles I_j quelconques vérifient ces propriétés. Désignons par :

$$G(s,t) = \sum_{n=0}^{\infty} s^n\, p_n(t)$$

la fonction génératrice des $p_n(t)$. D'après a/ et b/, $N(t_1 + t_2)$ est la somme de deux variables indépendantes admettant mêmes lois que $N(t_1)$ et $N(t_2)$ respectivement, d'où :

$$(6) \qquad G(s, t_1 + t_2) = G(s,t_1)\, G(s,t_2)$$

(équation des demi-groupes). D'après c/, d'autre part, on a :

$$G(s, \delta t) = 1 - a(1 - s)\,\delta t + O(\delta t)$$

D'où

D'où :

$$G(s,\delta t) - G(s,t) = - a(1-s)\, G(s,t)\, \delta t + O(\delta t)$$

Par suite, $G(s,t)$ admet la dérivée à droite :

$$\frac{\partial^+}{\partial t} G(s,t) = - a(1-s)\, G(s,t)$$

Il suffit de réecrire (6) sous la forme :

$$G(s, t-\delta t) = \frac{G(s,t)}{G(s,\delta t)} = G(s,t)\left[1 + a(1-s)\delta t\right] + O(\delta t)$$

pour voir que la dérivée à gauche existe aussi pour $t > 0$ et vaut encore $- a(1-s)G$. On a donc :

$$\frac{\partial}{\partial t} G(s,t) = - a(1-s)\, G(s,t)$$

D'où :

$$G(s,t) = C(s)\, e^{-a(1-s)t}$$

et $C = 1$, puisque $G(s,t) \to 1$ pour $t \to 0$.

Ainsi, $N(t)$ obéit bien à la loi de Poisson de paramètre at. L'équivalence $N(t) < n \Leftrightarrow T_n > t$ montre ensuite, selon un calcul déjà fait :

$$P(T_n > t) = \sum_{k=0}^{n-1} \frac{(at)^k}{k!} e^{-at} = \int_t^{\infty} \frac{a^n t^{n-1}}{(n-1)!} e^{-at}\, dt$$

et T_n obéit donc à la loi gamma de densité $\frac{a^n t^{n-1}}{(n-1)!} e^{-at}$

Remarque. La démonstration précédente n'est pas complète. On a seulement montré que $T_n = S_1 + \ldots + S_n$ obéit à la même loi que la somme de n variables indépendantes de densité $a\, e^{-at}$, et que chacune des S_i admet bien cette loi. Il n'en résulte pas encore, en toute rigueur, que les S_i soient effectivement indépendantes.

Pour le montrer, on pourrait invoquer l'absence de mémoire b/ : si l'on place l'origine des temps en T_{n-1}, les évènements postérieurs à T_{n-1} (en particulier $S_n = T_n - T_{n-1}$, $S_{n+1}, \ldots$) sont indépendants des évènements antérieurs à T_{n-1} (en particulier de $S_1, S_2, \ldots S_{n-1}$)

Il y a toutefois une difficulté : la propriété b/ s'applique à des intervalles d'extrémités données (non aléatoires), et, en toute rigueur, il n'est pas évident qu'elle s'étend à des intervalles d'extrémités aléatoires tels que $(0, T_{n-1})$.

Pour tourner cette difficulté, nous allons déterminer directement la loi de la variable à n composantes $(S_1, S_2 \dots S_n)$: Soient, tout d'abord, des instants $t_1, t_1 + \delta t_1, \dots$ avec :

$$0 < t_1 < t_1 + \delta t_1 < \dots < t_n < t_n + \delta t_n$$

L'événement : $\{ t_1 \leq T_1 \leq t_1 + \delta t_1, \dots t_n \leq T_n \leq t_n + \delta t_n \}$ équivaut à :

$$\{ N(t_1) = 0,\ N(t_1 + \delta t_1) - N(t_1) = 1, \dots ,\ N(t_n) - N(t_{n-1} + \delta t_{n-1}) = 0,\ N(t_n + \delta t_n) - N(t_n) \dots$$

dont la probabilité, d'après a/, b/ et la loi de $N(t)$, est

$$e^{-a[t_n + \delta t_n]}\ a^n \delta t_1 \delta t_2 \dots \delta t_n$$

On en déduit que la variable vectorielle $(T_1, \dots T_n)$ admet la densité

$$f(t_1, t_2, \dots t_n) = a^n e^{-a t_n}$$

pour $0 \leq t_1 \leq \dots \leq t_n$ (et 0 ailleurs). Le changement de variables :

$$S_1 = T_1 , \quad S_2 = T_2 - T_1 , \quad \dots \quad S_n = T_n - T_{n-1}$$

montre ensuite que $(S_1, S_2, \dots S_n)$ admet la densité :

$$a^n e^{-a(s_1 + s_2 + \dots + s_n)} \qquad (s_1 \geq 0,\ s_2 \geq 0,\ \dots s_n \geq 0)$$

et par suite les S_n sont bien des variables exponentielles indépendantes.

EXERCICES

Exercice 1 - Processus de Poisson pris conditionnellement dans l'hypothèse où $N(t_0) = n$.

a/ Soit $0 < t < t_0$. Conditionnellement en $N(t_0) = n$, le nombre $N(t)$ obéit à la loi binomiale :

$$P(N(t)=k \mid N(t_0)=n) = C_n^k \left(\frac{t}{t_0}\right)^k \left(1-\frac{t}{t_0}\right)^{n-k} \qquad (0 \leq k \leq n)$$

b/ Plus généralement, si l'on implante n points au hasard indépendamment les uns des autres et avec une densité uniforme sur le segment $(0, t_0)$, et si l'on désigne par $X_1, X_2 .. X_n$ leurs abscisses rangées par ordre croissant, la loi de $(X_1, X_2, .. X_n)$ coïncide avec la loi de $(T_1, T_2, .. T_n)$, temps d'attente d'un processus de Poisson pris conditionnelle dans l'hypothèse $N(t_0) = n$. Cette loi commune admet la densité :

$$\frac{n!}{t_0^n} \, dx_1 \, dx_2 \ldots dx_n$$

pour $0 \leq x_1 \leq .. \leq x_n \leq t_0$, et 0 ailleurs.

Exercice 2 - Processus de Poisson composés, ou processus à multiplets poissoniens.

Dans la caractérisation du processus de Poisson (5º/ ci-dessus), conservons les conditions a/ et b/, et remplaçons c/ par :

$$\begin{cases} p_0(\delta t) = 1 - a\,\delta t + \varepsilon(\delta t) \\ p_n(\delta t) = a_n\,\delta t + \varepsilon_n(\delta t) \\ \sum_{n=1}^{\infty} a_n = a \end{cases}$$

avec des infiniment petits $\varepsilon(\delta t)$, $\varepsilon_n(\delta t)$ d'ordre supérieurs à 1, tels que $\sum |\varepsilon_n(\delta t)|$ soit lui-même d'ordre supérieur à 1. On posera :

$$\begin{cases} \varpi_n = \dfrac{a_n}{a} \qquad \left(\sum_{n=1}^{\infty} \varpi_n = 1\right) \\ \gamma(s) = \sum_{n=1}^{\infty} \varpi_n s^n \end{cases}$$

a/ Montrer que $p_0(t) = e^{-at}$ et que l'on a encore, pour la fonction génératrice $G(s,t)$ des $p_n(t)$, la relation des demi-groupes :

$$G(s, t + t') = G(s,t)\, G(s,t')$$

b/ En déduire que G admet une dérivée partielle en t vérifiant :

$$\frac{\partial G}{\partial t} = -a[1-\gamma]\,G$$

et que l'on a par suite : $G(s,t) = e^{-a[1-\gamma(s)]t}$

c/ Mettre $G(s,t)$ sous la forme :

$$G(s,t) = \prod_{n=1}^{\infty} e^{-a_n(1-s^n)t}$$

Interpréter : $N(t)$ est une somme de multiplets poissoniens indépendants. Par exemple, s'il s'agit de particules enregistrées par un compteur, les particules arrivent les unes individuellement, d'autres par paires, ou par triplets etc... S'il s'agit d'accidents d'automobiles, certains de ces accidents ne mettent en jeu qu'une seule voiture, d'autres n voitures à la fois etc...

Pour chaque n donné, les n-tuplets (les paquets de n particules arrivant simultanément au compteur, les accidents mettant en jeu exactement n autos) constituent un processus de Poisson de paramètre a_n indépendant de tous les autres n'-tuplets.

La probabilité pour qu'un n-tuplet (n quelconque) survienne pendant le temps dt est a dt, de sorte que le nombre total des n-tuplets (nombre des paquets de particules, et non plus nombre des particules elles-mêmes) constitue un processus de Poisson de paramètre a.

Enfin, conditionnellement lorsque l'on sait qu'un paquet de particules est arrivé au temps t, la probabilité pour que ce paquet soit un n-tuplet est $\varpi_n = a_n/a$.