Intervalles de temps et comptages dans les processus ponctuels

Bernard Picinbono

Laboratoire des Signaux et Systèmes
Supélec, Plateau de Moulon, 91190 Gif-sur-Yvette, France
bernard.picinbono@lss.supelec.fr

Résumé – La relation entre la loi de probabilité des intervalles et celle des comptages dans les processus ponctuels est connue dans le cas stationnaire mais son extension au cas non-stationnaire conduit à des calculs erronés dus au fait de ne pas avoir considéré les distances entre points comme des variables aléatoires conditionnelles. Partant de cette remarque, le calcul complet et exact est présenté aussi bien pour le temps de survie que pour celui de vie. Ce calcul qui retrouve évidemment les résultats connus du cas stationnaire est appliqué à divers cas particuliers en particulier ceux des processus de Poisson stricts ou composés.

Abstract – The relationship between the probability law of time intervals between points and of counting is well known in the stationary case. However its extension to the non-stationary case yields to erroneous calculations. This is due to the fact that distances between points are not considered as conditional random variables. Starting from this remark, the complete and exact calculation is presented. It is valid as well for the life or the residual times. This calculation yields of course the results known in the stationary case but also some new results especially for pure or compound non-stationary Poisson processes.

1 Introduction

Les processus ponctuels jouent un rôle important dans de nombreuses questions de traitement du signal. Ils apparaissent à l'échelle microscopique par exemple pour décrire l'effet de grenaille électronique ou le bruit en transmission optique de l'information, mais aussi à l'échelle macroscopique dans l'étude des problèmes de trafic et d'encombrement des systèmes de communications.

Il y a deux manières traditionnelles de les analyser. La première fait intervenir les *comptages* de points dans des intervalles, les systèmes physiques correspondants utilisant essentiellement des compteurs. La seconde fait intervenir les *intervalles* entre points et se réalise expérimentalement à l'aide de convertisseurs temps-amplitude. La relation entre comptages et intervalles est bien établie *dans le cas stationnaire* [1][2]. Par contre son extension au cas non stationnaire pose des problèmes qui sont le point de départ de cet exposé.

L'idée fondamentale consiste à noter que les intervalles de temps entre points d'un processus ponctuels doivent être considérés comme des variables aléatoires (VA) *conditionnelles.* En effet, alors que pour un processus stationnaire de densité non nulle il y a toujours des points postérieurs à n'importe quel instant, ceci n'a aucune raison d'être vrai dans le cas non-stationnaire. Dès lors tous les calculs de loi des intervalles doivent être conditionnés au fait que l'événement étudié peut se produire, c'est à dire qu'il y a des points postérieurs dont les intervalles les séparant peuvent être analysés.

Après avoir précisé cette idée pour le cas du temps de survie d'ordre 1 on présente le calcul complet des lois de probabilité des temps de survie et de vie exprimées en fonction de la loi des nombres de points dans des intervalles quelconques. Les résultats sont illustrés par divers exemples liés aux processus de Poisson stricts ou composés.

2 Origine du problème

Pour expliquer l'idée de base, commençons par l'exemple le plus simple du temps de survie dans les processus de Poisson non-stationnaires. Ce cas est étudié dans des ouvrages de base [3][4][5] qui introduisent un résultat erroné fondé sur le raisonnement suivant. Soit L la variable aléatoire (VA) égale à la distance entre un instant quelconque pris comme origine du temps et le premier point du processus postérieur à cet instant (temps de survie). Pour calculer la densité de probabilité (ddp) $p(l)$ de L on part de l'événement que l'intervalle $[0, l[$ ne contient aucun point du processus et que l'intervalle $[l, l + dl[$ en contient au moins 1. Ces intervalles étant disjoints, les nombres de points associés sont des VA indépendantes, ce qui est la propriété de base d'un processus de Poisson, et dès lors la ddp cherchée s'écrirait

$$f(l) = u(l) \exp\left\{ -\int_0^l \lambda(\theta) d\theta \right\} \lambda(l), \qquad (1)$$

où $\lambda(t)$ est la densité du processus et $u(.)$ la fonction échelon unité. Dans le cas stationnaire on retrouve la loi exponentielle $u(l)\lambda \exp(-\lambda l)$. Mais il est facile de vérifier que $f(l)$ donnée par (1) *n'est pas nécessairement normée* et il ne s'agit donc pas d'une vraie ddp. Bien entendu le même problème se pose dans le cas du temps de vie qui est la distance entre deux points successifs du processus. L'erreur du raisonnement apparaît aussi lorsque l'on traite de la même manière des processus qui ne sont pas de Poisson. Il faut donc reprendre l'ensemble du problème.

L'idée de base qui n'a jamais été remarquée est que la VA L doit être considérée comme une VA *conditionnelle.* En effet elle n'est définie que s'il existe au moins un point du processus ponctuel postérieur à 0. Cet événement est presque sûr dans le cas stationnaire, mais n'a aucune raison de l'être dans le cas non stationnaire, ce qui est l'origine de l'erreur. La fonction de répartition (fr) $F(l)$ de L est donc une probabilité conditionnelle définie par

$$F(l) \triangleq P[L \leq l] = P(\bar{A}|\bar{D}) = \frac{P(\bar{A}\bar{D})}{P(\bar{D})} , \qquad (2)$$

où $\bar{A}$ est l'événement qu'il y a au moins un point dans l'intervalle $[0, l[$ et $\bar{D}$ celui qu'il y a au moins un point postérieur à 0. Soit alors $p_i(l)$ la probabilité qu'il y ait i points dans $[0, l[$ (loi des comptages) et π_i celle qu'il y ait i points du processus postérieurs à 0. On a évidemment $\pi_i = p_i(+\infty)$ et $P(\bar{D}) = 1 - \pi_0$. De plus on a $P(\bar{A}\bar{D}) = P(\bar{A})P(\bar{D}|\bar{A})$. Mais il est évident que $P(\bar{D}|\bar{A}) = 1$. On a donc $P(\bar{A}\bar{D}) = P(\bar{A}) = 1 - p_0(l)$, ce qui donne $F(l) = [1-\pi_0]^{-1}[1-p_0(l)]$. Il est facile de vérifier qu'il s'agit bien d'une fr (fonction non décroissante variant de 0 à 1 quand l varie de 0 à l'infini), de sorte que la ddp cherchée vaut

$$f(l) = \frac{-1}{1-\pi_0}\,\frac{\mathrm{d}p_0(l)}{\mathrm{d}l} = \frac{-1}{1-p_0(+\infty)}\,\frac{\mathrm{d}p_0(l)}{\mathrm{d}l} , \qquad (3)$$

qui est toujours normée, contrairement à (1), et valable quel que soit le processus ponctuel considéré. Cette équation établit bien la relation cherchée entre la loi des comptages et celle du temps de survie. En fait seule la probabilité p_0 pour qu'un intervalle soit vide est utilisée.

Dans le cas d'un processus de Poisson non-stationnaire de densité $\lambda(t)$ la relation (3) s'écrit

$$f(l) = \frac{1}{1-\mathrm{e}^{-M}}\,\mathrm{e}^{-m(l)}\,\lambda(l), \qquad (4)$$

où $m(l) = \int_0^l \lambda(\theta)d\theta$ et $M = m(\infty)$. C'est ce dernier terme de normalisation qui est oublié dans (1).

3 Temps de survie d'ordre n

Reprenant la même idée dans le cas plus complexe de la VA L_n distance entre l'origine et le $n$ième point du processus ponctuel postérieur à 0 on obtient sa fr par la relation

$$F_n(l) \triangleq P[L_n \leq l] = P(\bar{A_n}|\bar{D_n}) = \frac{P(\bar{A_n}\bar{D_n})}{P(\bar{D_n})} , \qquad (5)$$

où $\bar{A_n}$ est l'événement qu'il y a au moins n points dans l'intervalle $[0, l[$ et $\bar{D_n}$ celui qu'il y a au moins n points postérieurs à 0. Comme précédemment on a $P(\bar{A_n}\bar{D_n}) = P(\bar{A_n})$ et de plus $P(\bar{A_n}) = 1 - \sum_{i=0}^{n-1} p_i(l)$. Compte-tenu d'une relation similaire pour $P(\bar{D_n})$, on obtient

$$F_n(l) = \frac{1}{1-\sum_{i=0}^{n-1}\pi_i}\left[1 - \sum_{i=0}^{n-1} p_i(l)\right]. \qquad (6)$$

La ddp s'obtient par dérivation, soit

$$f_n(l) = \frac{-1}{1-\sum_{i=0}^{n-1} p_i(\infty)} \sum_{i=0}^{n-1} \frac{\mathrm{d}p_i(l)}{\mathrm{d}l}. \qquad (7)$$

Cette relation montre que pour calculer la ddp du temps de survie d'ordre n il faut connaître les lois des comptages de 0 à $n-1$ points.

Dans le cas d'un processus de Poisson non stationnaire de densité $\lambda(t)$ cette relation peut se mettre sous la forme

$$f_n(l) = \frac{1}{1-d_n(M)}\,\mathrm{e}^{-m(l)}\,\frac{m^{n-1}(l)}{(n-1)!}\,\lambda(l), \qquad (8)$$

où $m(l)$ et M ont été définis après (4) et

$$d_n(m) \triangleq \sum_{i=0}^{n-1} p_i = \mathrm{e}^{-m}\left[1 + m + ... + \frac{m^{n-1}}{(n-1)!}\right]. \qquad (9)$$

C'est ce dernier terme, généralement oublié, qui assure la normalisation de $f_n(l)$.

Considérons maintenant le cas des processus de Poisson composés. Ce sont des processus de Poisson dont la densité $\lambda(t)$ est aléatoire. Ils jouent un rôle fondamental en optique statistique [6]. La ddp du temps de vie d'ordre n est calculée à partir de (8) où $d_n(M)$ est remplacé par son espérance mathématique et où l'on prend l'espérance par rapport à $\lambda(t)$ du numérateur, ce qui donne

$$f_n(l) = \frac{1}{(n-1)!}\,\frac{\mathsf{E}\{[m(l)]^{n-1}\lambda(l)\mathrm{e}^{-m(l)}\}}{1-\mathsf{E}[d_n(M)]} . \qquad (10)$$

Dans le cas où $\lambda(t)$ est stationnaire on a $E[d_n(M)] = 0$, et l'on retrouve alors l'expression classique valable pour les processus de Poisson composés stationnaires.

4 Temps de vie

C'est la distance entre deux points successifs ou non du processus. Comme ci-dessus cette distance doit être considérée comme une VA conditionnelle dans le même sens que pour le temps de survie. Mais l'existence d'un point du processus à l'origine de l'intervalle complique sensiblement les calculs et l'on présente tout d'abord leur principe dans le cas le plus simple du temps de vie d'ordre 1.

Le raisonnement aboutissant à (3) est encore valable, à condition de remplacer p_0 et π_0 par des probabilités conditionnelles à l'existence d'un point du processus à l'origine dénommée ici t puisqu'il n'est plus possible de la fixer à l'avance. Il convient de noter que pour un processus *régulier* la probabilité qu'il y ait un point en un instant donné est nulle. Il faut donc considérer un intervalle infinitésimal au voisinage de t. Soit alors $p_0(t, l)$, $t < l$, la probabilité pour qu'il n'y ait aucun point du processus dans l'intervalle $[t, l[$, $t < l$. On fait une partition de cet intervalle en posant $I = [t, t + \Delta t[$ et $A = [t + \Delta t, l[$. Soit A_0 et I_0 respectivement les événements que A et I ne contiennent aucun point. Comme Δt est très petit, I ne peut contenir que 0 ou 1 point, ce dernier événement étant noté I_1. De plus $P(I_1) = \mu(t)\Delta t$ où $\mu(t)$ est la densité du processus ponctuel. La probabilité conditionnelle $\hat{p}_0(t, l)$ qui doit être insérée dans (3) est $\hat{p}_0(t, l) = P(A_0|I_1) = P(A_0 I_1)/P(I_1)$. Mais comme I_1 est le complémentaire de I_0 on a

$$P(A_0 I_1) = P(A_0) - P(A_0 I_0) = \frac{\partial p_0(t, l)}{\partial t} \Delta t. \qquad (11)$$

En faisant le même calcul pour $\hat{\pi}_0 = \hat{p}_i(t, +\infty)$ on obtient la fr du temps de vie par

$$g_1(t, l) = \frac{\mu(t) - \frac{\partial}{\partial t} p_0(t, l)}{\mu(t) - \frac{\partial}{\partial t} p_0(t, \infty)} \quad . \qquad (12)$$

On en déduit la ddp par dérivation, soit

$$g_1(t, l) = \frac{-1}{\mu(t) - \frac{\partial}{\partial t} p_0(t, \infty)} \, \frac{\partial^2 p_0(t, l)}{\partial t \partial l} \, . \qquad (13)$$

Il est facile de voir que cette quantité est bien normée lorsque l varie de t à l'infini. Il importe de bien voir la signification de cette relation. Il s'agit de la ddp d'une VA $L(t)$ distance entre un point du processus à l'instant t et le premier point postérieur à t. Il est évident que cette ddp dépend de t et de l. Mais il convient de remarquer que, comme pour le temps de survie, la loi de probabilité du temps de vie ne dépend que de la probabilité $p_0(t, l)$ pour que l'intervalle $[t, l[$ soit vide.

Le calcul peut se généraliser à l'ordre n mais nécessite quelques précautions. La fr $G_n(t, l)$ se déduit de (6) en remplaçant les probabilités marginales p_i par les probabilités conditionnelles $\hat{p}_i$. Comme pour $\hat{p}_0(t, l)$ on peut écrire $\hat{p}_i(t, l) = P(A_i|I_1) = P(A_i I_1)/P(I_1)$. Le dénominateur s'écrit toujours $P(I_1) = \mu(t)\Delta t$ et il reste donc à calculer le numérateur. Pour ceci on part du fait que l'intervalle I ne peut contenir que 0 ou 1 point, ce qui donne

$$P(A_i I_1) = P(A_i) - P(A_i I_0). \qquad (14)$$

Soit alors H l'intervalle $[t, l[= I \cup A$. Il est alors clair que l'événement H_i peut se décomposer en la somme $H_i = A_i I_0 + A_{i-1} I_1$ d'où il résulte

$$P(H_i) = P(A_i I_0) + P(A_{i-1} I_1). \qquad (15)$$

Combiné avec la relation précédente, ceci donne

$$P(A_i I_1) = P(A_i) - P(H_i) + P(A_{i-1} I_1). \qquad (16)$$

On peut répéter cette opération pour arriver à la relation (11) $P(A_0 I_1) = P(A_0) - P(H_0)$. Ceci donne donc

$$P(A_i I_1) = \sum_{k=0}^{i} [P(A_i) - P(H_i)]. \qquad (17)$$

Mais il résulte de la définition de H_i que

$$P(A_i) - P(H_i) = \frac{\partial p_0(t, l)}{\partial t} \Delta t. \qquad (18)$$

En insérant ces relations dans (6) on obtient

$$G_n(t, l) = \frac{\mu(t) - \sum_{i=0}^{n-1} \sum_{k=0}^{i} \frac{\partial}{\partial t} p_o(t, l)}{\mu(t) - \sum_{i=0}^{n-1} \sum_{k=0}^{i} \frac{\partial}{\partial t} p_o(t, \infty)} \, . \qquad (19)$$

En prenant la dérivée par rapport à l on obtient la ddp qui vaut

$$g_n(t, l) = \frac{-1}{D} \sum_{i=0}^{n-1} \sum_{k=0}^{i} \frac{\partial^2 p_k(t, l)}{\partial t \partial l} \, . \qquad (20)$$

où D est le dénominateur apparaissant dans (19).

Cette formule établit bien la relation entre la loi des comptages caractérisée par les $p_i(t, l)$ et celle de la distance entre un point du processus et le $n$ième qui lui est postérieur.

Malgré son apparence complexe elle donne pour les processus de Poisson purs ou composés non stationnaires des résultats très maniables.

On peut d'ailleurs la simplifier en notant que

$$\sum_{i=0}^{n-1} \sum_{k=0}^{i} \frac{\partial^2 p_k(t, l)}{\partial t \partial l} = \sum_{k=0}^{n-1} (n - k) \frac{\partial^2 p_k(t, l)}{\partial t \partial l} \, . \qquad (21)$$

5 Exemples

5.1 Processus de Poisson purs non-stationnaires

Soit un processus de Poisson dont la densité vaut λ dans l'intervalle $[0, T[$ et 0 ailleurs. La ddp de la variable aléatoire L_n, temps de survie d'ordre n est donnée par (8) qui devient

$$f_n(l) = \frac{1}{1 - d_n(M)} \, \frac{(\lambda l)^{n-1}}{(n-1)!} \, \lambda e^{-\lambda l} \, , 0 \le l \le T, \qquad (22)$$

et 0 ailleurs. On peut aisément calculer la moyenne de L_n qui vaut

$$\mathsf{E}(L_n) = \frac{n}{\lambda} \, \frac{1 - d_{n+1}(M)}{1 - d_n(M)} \, = \frac{n}{\lambda} \, c_n(M), \qquad (23)$$

où d_n and M ont été définis précédemment. Lorsque $T \to \infty$, $c_n(M) \to 1$ et le processus tend à devenir stationnaire. Dans ce cas $\mathsf{E}(L_n)$ tend vers la valeur n/λ, ce qui est bien en accord avec le fait que pour un processus de Poisson stationnaire de densité λ les distances entre points successifs sont des variables aléatoires IID de valeur moyenne $1/\lambda$. Ainsi le dernier terme $c_n(M)$ peut être considéré comme un facteur de correction lié à la non-stationnarité du processus. Il est représenté sur la figure 1 en fonction de M pour diverses valeurs de n. Il est clair que le facteur de normalisation $1 - d_n(M)$ est fondamental pour obtenir la vraie valeur de la moyenne.

La différence entre les processus de Poisson stationnaires ou non à densité constante est particulièrement importante quand la densité λ devient très petite. Dans le cas stationnaire la moyenne $1/\lambda$ tend vers l'infini. Ceci est évidemment impossible dans le cas étudié ici où le processus est limité à l'intervalle fini $[0, T[$. Pour $\lambda \to 0$, la moyenne $\mathsf{E}(L_n)$ définie par (23) tend vers $T[n/(n+1)]$.

Par exemple pour $n = 1$ ceci donne $T/2$, qui peut aussi s'obtenir à partir de le ddp $f_1(l)$ qui tend vers $1/T$ indiquant que L_n est uniformément distribué dans l'intervalle $[0, T[$.

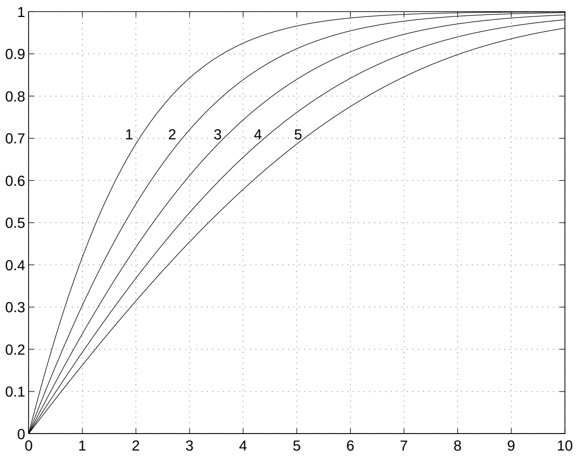

Figure 1. Terme de correction $c_n(M)$ en fonction de M pour $n = 1, 2, 3, 4, 5$.

5.2 Processus de Poisson composés

La loi du temps de survie est donnée par (10). Intéressons-nous donc à celle du temps de vie. Pour l'ordre 1 (intervalles entre points successifs) on part de la formule générale (12). Comme $p_0(t, l) = \mathsf{E}\{\exp[-m(t, l)]\}$ où $m(t, l)$ est la VA définie par

$$m(t, l) = \int_t^l \lambda(\theta) d\theta, \tag{24}$$

on obtient par dérivation de cette fonction

$$g_1(t, l) = \frac{1}{\mu(t) - \mathsf{E}[\lambda(t)\mathrm{e}^{-M}]} \; \mathsf{E}[\lambda(t)\lambda(l)\mathrm{e}^{-m(t,l)}]. \tag{25}$$

Dans le cas stationnaire $\mu(t) = \mu$ et $\mathsf{E}[\lambda(t)\mathrm{e}^{-M}] = 0$. On retrouve alors une formule classique. Les termes $\lambda(t)$ et $\lambda(l)$ sont reliés à l'existence d'un point aux voisinages de t et l.

Dans le cas général, un calcul non reproduit ici donne le résultat suivant

$$g_n(t, l) = \frac{1}{\mu(t) - \mathsf{E}[\lambda(t)\mathrm{e}^{-M}]} \; \mathsf{E}\left[\lambda(t)\frac{m(t, l)^{n-1}}{(n-1)!} \; \lambda(l)\mathrm{e}^{-m(t,l)}\right] \tag{26}$$

où $m(t, l)$ est la VA définie par (24). Cette formule, bien que plus compliquée à obtenir que (25) s'interprète de manière identique.

5.3 Processus stationnaires

Dans ce cas $T \to \infty$ et $\mu(t) = \mu$. Le dénominateur de (20) devient μ et compte tenu de (21) on obtient

$$g_n(t, l) = \frac{-1}{\mu} \sum_{k=0}^{n-1} (n-k) \frac{\partial^2 p_k(t, l)}{\partial t \partial l} \, . \tag{27}$$

Mais dans le cas stationnaire $p_k(t, l)$ ne dépend que de $l - t$, soit $f_k(l - t)$. Ceci donne une autre forme de (27) s'écrivant

$$g_n(t, l) = \frac{1}{\mu} \sum_{k=0}^{n-1} (n-k) f_k''(l - t). \tag{28}$$

On retrouve ainsi la relation (81) apparaissant dans [1]. Toutefois l'usage de (27) et (28) requiert une certaine attention et, pour simplifier, ceci va être analysé pour la ddp $g_1(t, l)$ d'un processus de Poisson composé. Pour un tel processus on a

$$p_0(t, l) = \mathsf{E}\left\{\exp[-\int_t^l \lambda(\theta) d\theta]\right\} \tag{29}$$

qui est effectivement une fonction de $l - t$ seulement parce que $\lambda(t)$ est un signal aléatoire stationnaire. L'application de (27) donne

$$g_1(t, l) = \frac{1}{\mu} \, \mathsf{E}\left\{\lambda(t)\lambda(l) \exp[-\int_t^l \lambda(\theta) d\theta]\right\}, \tag{30}$$

où $\mu = \mathsf{E}[\lambda(t)]$. Cette relation peut d'ailleurs être obtenue directement à partir des propriétés des processus de Poisson composés (voir p. 181 de [5]).

Par contre l'application de (28) donne

$$g_1(t, l) = \frac{1}{\mu} \, \mathsf{E}\left\{[-\lambda'(l) + \lambda^2(l)] \exp[-\int_t^l \lambda(\theta) d\theta]\right\}. \tag{31}$$

Il n'est pas évident de montrer l'identité de ces deux expressions. Cependant il est clair que (30) est beaucoup plus facile d'usage car on n'utilise pas la dérivée de la fonction aléatoire $\lambda(t)$. Mais son interprétation est également plus claire car, comme on l'a déjà noté, $\lambda(t)$ et $\lambda(l)$ sont directement liés à la présence d'un point en t et l, tandis que le terme exponentiel indique l'absence de points entre ces deux instants.

Références

[1] J. A. McFadden, "The axis crossing intervals of random functions," *IRE Trans. Inf. Theory*, Vol IT4, pp. 14-24, 1958.

[2] J. A. McFadden, "On the lengths of intervals in a stationary point process," *Journ. of Royal Stat. Soc.*, B 24, pp. 364-382, 1962.

[3] D. R. Cox and V. Isham, *Point processes.* London : Chapman and Hall, 1980. Prentice Hall, 1966.

[4] D. L. Snyder and M. I. Miller, *Random point processes in time and space.* New-York : Springer Verlag, 1991.

[5] B. Picinbono, *Signaux aléatoires, tome 2, Fonctions aléatoires et modèles.* Paris : Dunod, 1995.

[6] B. Saleh and M. Teich, *Introduction to photon communication.* New-York : Wiley, 1991.