

A. BEHTANE

Les processus de surveillance. Étude et simulation d'un cas

RAIRO. Recherche opérationnelle, tome 15, n° 3 (1981), p. 279-286

http://www.numdam.org/item?id=RO_1981__15_3_279_0

© AFCET, 1981, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

LES PROCESSUS DE SURVEILLANCE. ÉTUDE ET SIMULATION D'UN CAS (*)

par A. BEHTANE ⁽¹⁾

Résumé. — Il s'agit pour un observateur de surveiller et d'observer un événement qui arrive à des instants aléatoires.

Un modèle mathématique donne le nombre aléatoire et le nombre moyen d'événements observés à un instant donné, compte tenu de l'évolution de l'état de l'observateur et de la loi d'arrivée de l'événement.

Un certain nombre de simulations ont été réalisées, en modifiant les paramètres caractérisant le comportement de l'état de l'observateur, en prenant diverses lois d'arrivée de l'événement (lois : de Poisson, d'Erlang, de Weibull, exponentielle tronquée, hyperexponentielle) et enfin avec des événements de types différents.

Ces simulations ont permis d'avoir des résultats qui sont plus difficiles à obtenir par des calculs.

Mots clés : Modèle probabiliste, surveillance, simulation, système.

Abstract. — An observer has to supervise and observe an event that occurs at random instants.

A mathematical pattern gives the random number and the mean number of the observed events at a given time, depending on the evolution of the observer's state and on the distribution function of the occurrence of the event.

Some simulations are realised, by changing the parameters characterising the behaviour of the observer's state and by taking various distribution functions of the occurrence of the event (Poisson, Erlang, Weibull, truncated exponential, hyperexponential) and finally by using events of different types.

These simulations give results that are difficult to obtain by calculation.

Key words: Probabilistic models, monitoring, simulation, systems.

1. INTRODUCTION

Dans la classe des processus de service dont la définition et une formulation mathématique précises ont été données par M. le professeur R. FORTET, se trouvent les processus de surveillance que l'on rencontre notamment dans le comptage, la surveillance de processus industriels, d'engins aériens, de zones déterminées...

(*) Reçu avril 1980.

(1) Université de Paris-VI, Laboratoire de Calcul des Probabilités, 4, place Jussieu, 75005 Paris.

On remarque que toute information continue peut, presque toujours, se ramener à une information discrète.

On suppose donc que des événements arrivent dans l'intervalle de temps $[0, +\infty)$, suivant un processus de Poisson non homogène, de paramètre $r(t)$, et on pose :

$$R(t) = \int_0^t r(u) du \quad \text{avec} \quad R(0) = 0.$$

Un observateur, physique ou humain, \mathcal{O} , a pour mission de surveiller et d'observer ces événements. On suppose qu'il peut se trouver dans deux états possibles : un état attentif, A ou un état distrait, D , et qu'il a une probabilité d'observer un événement, qui est égale à p s'il se trouve dans l'état A et à q s'il se trouve dans l'état D . On suppose aussi que \mathcal{O} se trouve dans l'état A à l'instant initial et qu'il se remet dans cet état ou y reste s'il s'y trouve déjà, chaque fois qu'il observe un événement (sensibilisation). Il reste dans l'état A pendant une durée aléatoire obéissant à une loi de fonction de répartition $G(t)$; s'il quitte l'état A , il se met dans l'état D pendant une autre durée aléatoire obéissant à une loi de fonction de répartition $\bar{H}(t)$.

On note :

τ_k , l'instant d'observation du k -ième événement ($k = 1, 2, \dots$);

$N(t)$, le nombre aléatoire d'événements observés dans l'intervalle $[0, t]$;

$m(\alpha, t)$, le nombre moyen d'événements observés entre α et t , avec :

$$m(0, t) = m(t),$$

$P_k(\alpha, t)$, la probabilité que k événements sont observés entre α et t , avec :

$$P_k(0, t) = P_k(t),$$

$F_k(\alpha, t)$, la probabilité que k événements au moins sont observés entre α et t , avec :

$$F_k(0, t) = F_k(t); \quad F_1(\alpha, t) = F(\alpha, t); \quad F(\alpha, \alpha) = 0;$$

$P_A(t)$, la probabilité que \mathcal{O} est dans l'état A à l'instant t .

2. ÉTUDE FONCTIONNELLE

Le processus d'arrivée étant poissonien et $P_A(t)$ se renouvelant à chaque instant tel que τ_k ($k = 1, 2, \dots$) il en résulte que $\tau_{k+1} - \tau_k$ est une variable aléatoire ne dépendant que de τ_k :

$$\tau_{k+1} = \tau_1 + (\tau_2 - \tau_1) + \dots + (\tau_{k+1} - \tau_k),$$

c'est la somme de $k + 1$ variables aléatoires indépendantes mais n'ayant pas même fonction de répartition.

On a :

$$\begin{aligned} \Pr[\tau_{k+1} - \tau_k < x] &= 1 - \sum_{k=0}^{\infty} \Pr[k \text{ événements arrivés et non observés dans }]\tau_k, \tau_k + x[\\ &= 1 - \sum_{k=0}^{\infty} P_1 P_2 = 1 - \sum_{k=0}^{\infty} P_1 P_3^k, \end{aligned}$$

où :

- $P_1 = \Pr[k \text{ événements arrivés, dans }]\tau_k, \tau_k + x[;$
- $P_2 = \Pr[k \text{ événements sont observés sachant qu'ils sont arrivés, dans }]\tau_k, \tau_k + x[;$
- $P_3 = \Pr[1 \text{ événement supposé arrivé, n'est pas observé, dans }]\tau_k, \tau_k + x[.$

En effet $P_2 = P_3^k$ d'après une propriété bien connue du processus de Poisson à savoir, si dans un intervalle (t_1, t_2) l'effectif des événements arrivant suivant un processus de Poisson est connu, tout se passe comme si ces événements arrivaient indépendamment les uns des autres et suivant une même loi de fonction de répartition $(R(t) - R(t_1))/(R(t_2) - R(t_1))$. Alors on peut écrire :

$$P_1 = e^{-[R(\tau_k+x) - R(\tau_k)]} \times \frac{[R(\tau_k+x) - R(\tau_k)]^k}{k!},$$

$$P_3 = \int_{\tau_k}^{\tau_k+x} [(1-p)P_A(u - \tau_k) + (1-q)[1 - P_A(u - \tau_k)]] \times \frac{r(u)}{R(\tau_k+x) - R(\tau_k)} du,$$

$$P[\tau_{k+1} - \tau_k < x] = 1 - e^{-q[R(\tau_k+x) - R(\tau_k)] - (p-q) \int_{\tau_k}^{\tau_k+x} P_A(u - \tau_k) r(u) du}.$$

Il faut calculer $P_A(u - \tau_k)$ pour $u \in]\tau_k, \tau_k + x]$, c'est-à-dire $P_A(t)$ pour $t \in [0, x]$, en fonction de $G(t)$ et $H(t)$. En remarquant que le changement d'état de D en A est un processus de renouvellement, et en désignant par $P_k^1(y)$ la probabilité d'avoir au moins k changements d'état entre 0 et y et par $v_1(y)$ le nombre moyen de changements d'état entre 0 et y , on peut alors écrire :

$$v_1(y) = \sum_{k=1}^{\infty} P_k^1(y)$$

et :

$$P_A(t) = 1 - G(t) + \sum_{k=1}^{\infty} \int_0^t [1 - G(t-y)] d_y P_k^1(y),$$

soit :

$$P_A(t) = 1 - G(t) + \int_0^t [1 - G(t-y)] dv_1(y).$$

En particulier si $G(t)$ et $H(t)$ admettent des densités respectives $g(t)$ et $h(t)$ dont les transformées de Laplace respectives sont $g(s)$ et $h(s)$ et si \mathcal{L} désigne l'opérateur de la transformée de Laplace, on a :

$$\mathcal{L}[P_A(t)] = \frac{1-g(s)}{s[1-g(s)h(s)]} \Rightarrow P_A(t) = \mathcal{L}^{-1} \left[\frac{1-g(s)}{s[1-g(s)h(s)]} \right].$$

La formule précédente s'écrit aussi :

$$F(\alpha, t) = \Pr[\tau_{k+1} < t/\tau_k = \alpha] = 1 - e^{-g[R(t)-R(\alpha)] - (p-q) \int_{\alpha}^t P_A(u-\alpha)r(u)du}$$

d'où :

$$F_{k+1}(t) = \int_0^t F(\alpha, t) d_{\alpha} F_k(\alpha) = \int_0^t F_k(\alpha, t) dF(\alpha).$$

Comme :

$$P_k(t) = F_k(t) - F_{k+1}(t) \quad \text{avec} \quad F_0(0) = 0 \quad \text{et} \quad F_0(t) = 1,$$

alors :

$$P_k(t) = \int_0^t [1 - F(\alpha, t)] d_{\alpha} F_k(\alpha) \quad \text{et} \quad m(t) = \sum_{k=1}^{\infty} F_k(t),$$

$$m(\alpha, t) = \sum_{k=1}^{\infty} F_k(\alpha, t) = F(\alpha, t) + \int_{\alpha}^t \sum_{k=1}^{\infty} [F_k(u, t)] d_{\alpha} F(\alpha, u),$$

$$m(\alpha, t) = \int_{\alpha}^t [1 + m(u, t)] d_{\alpha} F(\alpha, u)$$

et, en particulier :

$$m(t) = \int_0^t [1 + m(u, t)] dF(u).$$

Ces formules permettent de calculer numériquement $F_k(t)$, $P_k(t)$ et $m(t)$.

Cas particuliers

Le processus d'arrivée est un processus de Poisson homogène de paramètre λ .
 $G(t)$ et $H(t)$ sont des exponentielles, avec $G(t) = 1 - e^{-at}$ et $H(t) = 1 - e^{-bt}$.

Alors :

$$\mathcal{L}[P_A(t)] = \frac{s+b}{s(s+a+b)} \Rightarrow P_A(t) = \frac{b}{a+b} + \frac{a}{a+b} e^{-(a+b)t},$$

$$F(t) = \Pr[\tau_{k+1} - \tau_k < t] = 1 - e^{-\lambda(aq+bp)/(a+b)t - (\lambda a(p-q)/(a+b)^2)(1-e^{-(a+b)t})},$$

d'où :

$$m(t) = \frac{\lambda(aq+bp)}{a+b} t + \frac{\lambda a(p-q)}{(a+b)^2} [1 - e^{-(a+b)t}],$$

pour t assez grand :

$$m(t) \simeq \frac{\lambda a(p-q)}{(a+b)^2} + \frac{\lambda(aq+bp)}{a+b} t.$$

3. SIMULATION

C'est le dernier cas particulier qui a d'abord été simulé.

Le programme de simulation permet de sortir :

- l'instant d'arrivée d'un événement : $X_k = t_k = \tau_k$;
- le nombre d'événements arrivés jusqu'à cet instant;
- l'état de l'observateur à cet instant;
- le nombre d'événements observés jusqu'à cet instant : $Y_k = N(t_k)$.

La simulation s'arrête lorsque ce dernier nombre atteint un nombre N fixé à l'avance et les résultats suivants sont imprimés :

- nombre d'événements arrivés Na ;
- nombre d'événements observés N ;
- équation de la courbe de régression : $v(t) = \hat{\alpha} + \hat{\beta} t = \hat{Y}(t)$;
- équation de la courbe théorique : $m(t) = \alpha_1 + \beta_1 t$;
- coefficient de corrélation ρ et coefficient de Snedecor F ;
- écart-type et intervalle de confiance sur α et β .

Rappelons les résultats du calcul de ces diverses grandeurs :

$$X_1 = \sum_{i=1}^N X_i, \quad Y_1 = \sum_{i=1}^N Y_i,$$

$$P_1 = \sum_{i=1}^N X_i Y_i, \quad X_2 = \sum_{i=1}^N X_i^2,$$

$\bar{X} = X_1/N$: moyenne des X_i , $\bar{Y} = Y_1/N$, moyenne des Y_i ;

$$S_1 = \sum_{i=1}^N (Y_i - \bar{Y})^2,$$

$$E_1 = \sum_{i=1}^N (\hat{Y}_i - \bar{Y})^2, \quad R_1 = \sum_{i=1}^N (Y_i - \hat{Y}_i)^2,$$

$S = S_1 / (N - 1)$: erreur quadratique moyenne totale;

$E = E_1 / (N - 2)$: erreur quadratique moyenne résiduelle;

$$\hat{\beta} = \frac{NP_1 - X_1 Y_1}{NX_2 - X_1^2}, \quad \hat{\alpha} = \frac{Y_1 - \hat{\beta} X_1}{N} = \bar{Y} - \hat{\beta} \bar{X},$$

$$\rho = \left(\frac{R_1}{S_1} \right)^{1/2}, \quad F = \frac{R_1}{E},$$

$$\sigma^2(B) = \frac{NE}{NX_2 - X_1^2}, \quad \sigma^2(\alpha) = \frac{EX_2}{NX_2 - X_1^2},$$

$$\sigma^2(Y_N) = E \cdot \left[1 + \frac{1}{N} + \frac{N(X_N - \bar{X})}{NX_2 - X_1^2} \right],$$

$$\Delta\beta = t(N - 2, P) \cdot \sigma(\beta), \quad \Delta\alpha = t(N - 2, P) \cdot \sigma(\alpha),$$

$$\Delta(Y_N) = t(N - 2, P) \cdot \sigma(Y_N),$$

où $t(N - 2, P)$ = coefficient de Student pour un intervalle de confiance à 100 P % est donné par le tableau des résultats et où $\Delta\beta, \Delta\alpha, \Delta(Y_N)$ sont les écarts sur β, α et Y_N .

Résultats de simulations

Trois simulations indépendantes réalisées avec les mêmes valeurs des paramètres :

$$L = \frac{1}{\lambda} = 50, \quad A = \frac{1}{a} = 50, \quad B = \frac{1}{b} = 5,$$

$$P = 0,8, \quad Q = 0,2,$$

ont donné les résultats suivants :

Résultats \ Simulations	1 ^{re} simulation	2 ^e simulation	3 ^e simulation
Nombre d'événements arrivés : Na	128	132	130
Nombre d'événements observés : N	100	100	100
Droite de régression : $v(t)$	1,992 + 0,0126 t	3,346 + 0,0167 t	6,622 + 0,0139 t
Droite théorique : $m(t)$	0,005 + 0,0149 t	0,005 + 0,0149 t	0,005 + 0,0149 t
Coefficient de corrélation : ρ	0,994	0,996	0,993
Coefficient de Snedecor : F	8459	11 129	6 566
Écart sur β : $\Delta\beta$ à 95 %	0,003	0,003	0,003
Écart sur α : $\Delta\alpha$ à 95 %	1,201	1,026	1,267
Instant d'observation du 100 ^e événement = t_{100}	7 200	5 852	6 771
$m(t_{100})$	107	87	101
$v(t_{100}) = \hat{Y}_{100}$	93	101	100
$\Delta v(t_{100})$ à 95 %	6	6	7
Intervalle confiance ($v - \Delta v, v + \Delta v$) à 95 %	(87,99)	(95,107)	(93,107)

On a noté : $m(t) = \alpha_1 + \beta_1 t$ et $v(t) = \alpha + \beta t$ (on peut utiliser α au lieu de $\hat{\alpha}$ et β au lieu de $\hat{\beta}$).

Les résultats d'une simulation prise toute seule, paraissent excellents lorsqu'on ne connaît pas l'équation de $m(t)$: ρ très élevé, $\Delta\beta$ très faible et Δv faible également.

Mais, d'une part, on voit que pour une même simulation, β est nettement différent de β_1 , d'autre part, en passant d'une simulation à une autre, β varie fortement : la dispersion sur β est très forte et $\Delta\beta$ calculé pour une seule simulation, n'a plus de signification d'ensemble.

Les résultats de la 3^e simulation sont les plus intéressants : β est voisin de β_1 , $v(t_{100})$ est très voisin de $m(t_{100})$.

Lorsqu'on sait que cette 3^e simulation est en fait une simulation qui a débuté depuis longtemps et dont on n'a prélevé qu'une partie correspondant à un régime permanent d'une part, et, ayant les mêmes caractéristiques que les deux autres pour pouvoir faire des comparaisons, d'autre part, on comprend l'importance du temps de simulation dans la convergence de la simulation.

On a également réalisé d'autres simulations dans lesquelles :

1° On fait varier P , les autres paramètres étant constants. On a obtenu : pour $P=0,8$, alors : $\beta=0,0167$; pour $P=0,95$, alors : $\beta=0,0176$.

Ces résultats sont en concordance avec la théorie :

2° On fait varier A et B , tout en gardant le rapport A/B constant, les autres paramètres étant constants aussi. On a obtenu pour $A=50$ et $B=5$: $\beta=0,0112$; et pour $A=100$ et $B=10$: $\beta=0,0108$. β reste dans son intervalle de confiance.

En effet :

$$\lambda \frac{aq + bp}{a + b} = \lambda \frac{q + (b/a)p}{1 + (b/a)}$$

ne dépend que de A/B et non des valeurs réelles de A et B , si λ , p et q sont constants.

3° On fait varier seulement λ (ou $L=1/\lambda$) pendant le même essai : d'après la courbe $N(t)$, l'observateur s'adapte rapidement au changement de fréquence d'arrivée d'événements, la pente de la courbe $N(t)$ change en effet dès que λ varie, dans des proportions conformes à la théorie; cependant le problème du transitoire demeure.

4° La loi d'arrivée est, non pas une loi de Poisson, mais une des lois suivantes : Erlang, Weibull, exponentielle tronquée, hyperexponentielle.

Par rapport à la loi de Poisson, les résultats obtenus sont plus réguliers et plus valables pour les lois d'Erlang et de Weibull, mais sont généralement plus

dispersés et plus mauvais pour les lois exponentielle tronquée et hyperexponentielle, comme on pouvait s'y attendre.

5° L'observateur surveille plusieurs types différents et indépendants d'événements $T(1), \dots, T(n)$; on s'aperçoit que, même dans le cas où chacun des types $T(i)$, ($i=1, \dots, n$) a la même probabilité $1/n$ d'être surveillé et où les périodes de surveillance de chaque type $T(i)$ obéissent à des lois de même fonction de répartition, les coefficients β_i et par conséquent les courbes $N_i(t)$, varient selon la valeur de i ; dans ce cas, le temps d'établissement du régime permanent augmente avec n et la simulation doit durer plus longtemps que précédemment.

4. CONCLUSION

Partant de certaines hypothèses, on a construit une théorie de la surveillance et on a trouvé un modèle mathématique de la surveillance; le phénomène de sensibilisation évoqué, constitue un caractère d'adaptation du système aux événements extérieurs alors que l'on a presque toujours supposé que les lois de fonctionnement des systèmes de service sont indépendantes des lois d'arrivée des clients (ou événements); cet aspect adaptatif augmente le rendement du système.

La validité du modèle dépend de l'analyse du processus, des hypothèses et de la formulation; il reste à être vérifié par des résultats expérimentaux. Si ces derniers ne peuvent être obtenus, on peut avoir recours à des simulations qui, bien qu'elles ne puissent se substituer à elles, s'en rapprochent davantage et permettent d'analyser statistiquement les résultats, d'estimer les valeurs les plus probables, les intervalles de confiance; la simulation permet aussi de faire varier les hypothèses de départ et voir les résultats correspondants et c'est ce qui a été fait lors des dernières simulations réalisées, les résultats étant concluants.

BIBLIOGRAPHIE

- R. FORTET, *Les répartitions ponctuelles markoviennes; estimations et tests concernant leurs lois de probabilités*, Annales de l'Institut Henri-Poincaré, vol. V, n° 3, 1969.
 D. R. COX, *The theory of Stochastic Processes*, Chapman and Hall Ltd., 1972.
 D. N. CHORAFAS, *Systems and Simulation*, Academic Press, 1965.
 A. ALBERT, *Regression and the Moore Penrose Pseudo Inverse*, Academic Press, 1972.
 A. BEHTANE, *Les modèles de surveillance*, Thèse de Docteur-Ingénieur, Paris, 1970.