Annales Univ. Sci. Budapest de R. Eötvös nom. Sectio Math. **2** (1959), pp. 139–146.

ON SECONDARY PROCESSES GENERATED BY RANDOM POINT DISTRIBUTIONS

András Prékopa (Budapest)
Mathematical Institute, Eötvös Loránd University, Budapest

Received: May 15, 1959

Introduction

In the paper [3] we have given a rigorous mathematical model and a general method for solving special problems in connection with secondary processes generated by a random point distribution of Poisson type. A random point distribution is a random selection of a finite or countably infinite number of points of an abstract set T where a σ-algebra $\mathcal{S}_T$ is given. It is supposed that if $\xi(A)$ denotes the number of random points lying in the set $A \in \mathcal{S}_T$ then it is a random variable i.e. $\xi(A)$ is measurable in the sample space (a sample element is a selection of a finite or countably infinite number of points of T). A random point distribution is called of Poisson type if to disjoint sets $A_1, \dots, A_r$ of $\mathcal{S}_T$ there correspond independent random variables $\xi(A_1), \dots, \xi(A_r)$ and there is a σ-finite measure $\lambda(A)$, $A \in \mathcal{S}_T$ such that

$$\mathbb{P}(\xi(A) = k) = \frac{\lambda^k(A)}{k!} e^{-\lambda(A)}, \qquad k = 0, 1, 2, \dots \tag{1}$$

for every finite $\lambda(A)$, $A \in \mathcal{S}_T$.

If every random point is the starting point of a further "happening" then we have a secondary "process". Let Y be an abstract space with a σ-algebra $\mathcal{S}_Y$ and suppose that the secondary happening consists of a random selection of an element $y \in Y$ to each random point $t \in T$. Then the whole phenomenon can be characterized as follows: we choose a finite or countable infinite number of points $t_1, t_2, \dots,$ from T, also the corresponding y's $y_1, y_2, \dots$ from Y and thus a sample element of the secondary process is a random selection of the points $(t_1, y_1), (t_2, y_2), \dots$ from the product space $T \times Y$. In [3] the case was considered where for different t's the corresponding y's are chosen independently.

If T is the positive part of the time axis and Y is the one-dimensional Euclidean space, then $T \times Y$ is the ensemble of points (t, y) where $t \geq 0$.

In this case if $(t_1, y_1), (t_2, y_2), \dots$ is a sample element of the whole phenomenon this can be represented on the plane so that we measure y_i on the line crossing at t_i the time axis and being vertical to it. This representation was applied by Lexis for the birth- and death-process.

In the present paper we generalize the results of [3] for the case when the parameter-measure of the underlying random point distribution given in T depends also on chance, and has the form $\lambda(A) = \gamma(A, \lambda)$ where λ is a random variable and $\gamma(A, \lambda)$ is a σ-finite measure on $\mathcal{S}_T$ for every fixed λ. In § 1 we formulate a theorem and in § 2 we illustrate the obtained results.

§ 1. Abstract mixed Poisson random point distributions and secondary processes

Definition. A random point distribution is called of mixed Poisson type if it is a mixture of random point distributions of Poisson type i.e. there is a random variable λ, a σ-finite non-atomic measure $\gamma(A, \lambda)$ $(A \in \mathcal{S}_T)$ (which is a function of λ) such that

$$\mathbb{P}(\xi(A_1) = k_1, \ldots, \xi(A_n) = k_n \mid \lambda) = \prod_{i=1}^{n} \mathbb{P}(\xi(A_i) = k_i \mid \lambda) = \prod_{i=1}^{n} \frac{\gamma^k(A_i, \lambda)}{k_i!} e^{-\gamma(A_i,\lambda)} \tag{2}$$

provided $A_i \in \mathcal{S}_T$, $\gamma(A_i, \lambda) < \infty$ $(i = 1, \ldots, n)$ and $A_i A_k = 0$ for $i \neq k$.

Before formulating our theorem we mention that we maintain the supposition about T introduced in [3] i.e. we suppose that

α) For every $B \in \mathcal{S}_T$ there is a sequence of decompositions $\{B_i^{(n)},\ i = 1, \ldots, r_n\}$ $B_i^{(n)} B_k^{(n)} = 0$ for $i \neq k$, $\sum_{i=1}^{n} B_i^{(n)} = B$, $B_i^{(n)} \in \mathcal{S}_T$ $(i = 1, \ldots, r_n)$ such that the sets $B_i^{(n)}$ can be decomposed by the sets $B_k^{(n+1)}$ and if $t_1 \in B$, $t_2 \in B$, $t_1 \neq t_2$ then for a sufficiently large n we have $t_1 \in B_i^{(n)}$, $t_2 \in B_k^{(n)}$ where $i \neq k$.

Let us consider the sample elements (t_1, y_1), $(t_2, y_2), \ldots$ of the secondary process the ensemble of which form a random point distribution in $T \times Y$. We denote by $\mathcal{S}_{T\times Y}$ the σ-algebra $\mathcal{S}_T \times \mathcal{S}_Y$ and by $\eta(B)$ $(B \in \mathcal{S}_{T\times Y})$ the number of random points lying in the set B which is now a measurable subset of $T \times Y$. The condition that the secondary happenings corresponding to different random points are independent is now replaced by the following one

β) If $t_1, t_2, \ldots$ is a sample element of the random point distribution given in the set T then for every n we have

$$\mathbb{P}(\eta(D_1) = 1, \ldots, \eta(D_n) = 1 \mid \lambda, t_1, t_2, \ldots) = \mu(C_1, \lambda, t_1), \ldots, \mu(C_n, \lambda, t_n), \tag{3}$$

where $D_i = A_i \times C_i$, $A_i \in \mathcal{S}_T$, $C_i \in \mathcal{S}_Y$, $t_i \in A_i$ $(i = 1, 2, \ldots)$ and $A_i A_k = 0$ for $i \neq k$.

Thus the independence is required for every fixed value of λ. $\mu(C, \lambda, t)$ denotes the probability distribution of the secondary happening y provided its starting point is $t \in T$ and λ is given. Now we are in the position to formulate our

Theorem. *If the random point distribution $\{(t_1, y_1), (t_2, y_2), \ldots\}$ is generated by the random point distribution of a mixed Poisson type $\{t_1, t_2, \ldots\}$ with the random parameter measure $\gamma(A, \lambda)$ $(A \in \mathcal{S}_T)$ moreover Conditions α) and β) hold, then the former one is also of mixed Poisson type and*

$$\mathbb{P}(\eta(D) = n) = \int_0^\infty \frac{(\nu^*(D, \lambda))^n}{n!} e^{-\nu^*(D,\lambda)}\, \mathrm{d}U(\lambda), \qquad n = 0, 1, 2, \ldots, \tag{4}$$

where ν^* is the extended measure of ν which is given for the rectangular sets $D = A \times C$ $(A \in \mathcal{S}_T,\ C \in \mathcal{S}_Y)$ by

$$\nu(D, \lambda) = \int_A \mu(C, \lambda, t)\gamma(\mathrm{d}t, \lambda) \tag{5}$$

when λ is fixed and $U(\lambda)$ is the probability distribution of the random variable λ.

Our theorem is a simple consequence of the results of [3] therefore we omit the proof.

§ 2. Illustration

A process of random events is called a mixed Poisson process[1] if

$$\mathbb{P}(\xi_t = n) = \int_0^\infty \frac{(t\lambda)^n}{n!} e^{-t\lambda}\, \mathrm{d}U(\lambda), \qquad n = 0, 1, 2, \dots, \tag{6}$$

where ξ_t is the number of events occurring during the time interval $(0, t)$ and $U(x)$ is a distribution function ($U(\lambda) = 0$ for $\lambda \le 0$) and if λ is fixed, then ξ_t has independent increments. More precisely, if $t_1 \le t_2 \le \dots \le t_{2n-1} \le t_{2n}$ then

$$\begin{aligned} &\mathbb{P}(\xi_{t_2} - \xi_{t_1} = k_1, \dots, \xi_{t_{2n}} - \xi_{t_{2n-1}} = k_n \mid \lambda) \\ &= \prod_{i=1}^n \mathbb{P}(\xi_{t_{2i}} - \xi_{t_{2i-1}} = k_i \mid \lambda) \\ &= \prod_{i=1}^n \frac{((t_{2i} - t_{2i-1})\lambda)^{k_i}}{k_i!} e^{-(t_{2i} - t_{2i-1})\lambda}. \end{aligned} \tag{7}$$

If $\xi(A)$ denotes the number of events or what is the same the number of random points lying in the linear Borel-set A then it follows from (6) and (7) that if $|A| < \infty$ where $|A|$ is the Lebesgue-measure of A, then

$$\mathbb{P}(\xi(A) = n \mid \lambda) = \frac{(|A|\lambda)^n}{n!} e^{-|A|\lambda}, \qquad n = 0, 1, 2, \dots \tag{8}$$

$$\mathbb{P}(\xi(A) = n) = \int_0^\infty \frac{(|A|\lambda)^n}{n!} e^{-|A|\lambda}\, \mathrm{d}U(\lambda), \qquad n = 0, 1, 2, \dots \tag{9}$$

moreover

$$\mathbb{P}(\xi(A_1) = k_1, \dots, \xi(A_n) = k_n \mid \lambda) = \prod_{i=1}^n \mathbb{P}(\xi(A_i) = k_i \mid \lambda), \tag{10}$$

where $A_1, \dots, A_n$ are disjoint linear Borel-sets with finite Lebesgue measures. These are simple consequences of the extension theory of stochastic set functions (see [2]). Thus the mixed Poisson process can be considered as a special random point distribution of mixed Poisson type.

[1] This process was considered first by O. Lundberg [1] who called it a compound Poisson process. Since the terminology "compound Poisson process" is often used for a process ξ_t with the characteristic function $\exp t \sum_{k=1}^{\infty} C_k(e^{iku} - 1)$ (called also a *Composed* Poisson process) we call the process with absolute probabilities (6) mixed Poisson process.

In the special case when $u(\lambda) = U'(\lambda)$ exists and

$$u(\lambda) = \frac{1}{b}\frac{\lambda^{\frac{1}{b}-1}}{\Gamma\left(\frac{1}{b}\right)}e^{-\frac{\lambda}{b}}, \qquad b > 0, \tag{11}$$

we obtain from (6)

$$\begin{aligned} P_n(t) &= \mathbb{P}(\xi_t = n) = \left(\frac{t}{1+bt}\right)^n \frac{(1+b)\cdots(1+(n-1)b)}{n!}P_0(t) \\ P_0(t) &= (1+bt)^{-\frac{1}{b}}. \end{aligned} \tag{12}$$

Such a process was called by LUNDBERG a Pólya process. Letting $b \to 0$, the probability distribution (12) tends to

$$\frac{t^n}{n!}e^{-t}, \qquad n = 0, 1, 2, \dots. \tag{13}$$

Let $p_n(t)$ denote the intensity function of the Pólya-process, i.e. $p_n(t)\Delta t + \sigma(\Delta t)$ is the probability of an event occurring during the interval $(t, t + \Delta t)$ under the condition that $\xi_t = n$, then a simple calculation shows that

$$p_n(t) = \frac{1+nb}{1+tb}, \qquad n = 0, 1, 2, \dots \tag{14}$$

(see [1]). It is not difficult to prove that

$$p_n(t) = \mathbb{M}(\lambda \mid \xi_t = n). \tag{15}$$

In fact, using the theorem of BAYES we obtain

$$\begin{aligned} \mathbb{M}(\lambda \mid \xi_t = n) &= \frac{\displaystyle\int_0^\infty \lambda\frac{(\lambda t)^n}{n!}e^{-\lambda t}u(\lambda)\,\mathrm{d}\lambda}{\displaystyle\int_0^\infty \frac{(\lambda t)^n}{n!}e^{-\lambda t}u(\lambda)\,\mathrm{d}\lambda} \\ &= \frac{\displaystyle\frac{n+1}{t}\int_0^\infty \frac{(\lambda t)^{n+1}}{(n+1)!}e^{-\lambda t}u(\lambda)\,\mathrm{d}\lambda}{\displaystyle\int_0^\infty \frac{(\lambda t)^n}{n!}e^{-\lambda t}u(\lambda)\,\mathrm{d}\lambda} = \frac{\frac{n+1}{t}P_{n+1}(t)}{P_n(t)} = \frac{1+nb}{1+tb}, \end{aligned} \tag{16}$$

where $u(\lambda)$ is the function given by (11) and $P_n(t)$ the probability given by (12).

If λ has the probability density

$$u(\lambda) = C^r\frac{\lambda^{r-1}}{\Gamma(r)}e^{-C\lambda}, \qquad r > 0, \quad C > 0 \tag{17}$$

which is a Pearson type III curve then instead of (11) we get

$$P_n(t) = \left(\frac{C}{C+t}\right)^r\left(\frac{t}{C+t}\right)^n\frac{r(r+1)\cdots(r+n-1)}{n!}, \qquad n = 0, 1, 2, \dots. \tag{18}$$

A simple argument shows that in this case

$$p_n(t) = \mathbb{M}(\lambda \mid \xi_t = n) = \frac{r+n}{C+t}, \tag{19}$$

$$\begin{aligned} \mathbb{M}(\xi_t) &= \sum_{n=1}^{\infty} nP_n(t) = \int_0^{\infty} \sum_{n=1}^{\infty} n\frac{(\lambda t)^n}{n!} e^{-\lambda t} u(\lambda)\,\mathrm{d}\lambda \\ &= t\int_0^{\infty} \lambda u(\lambda)\,\mathrm{d}\lambda = t\mathbb{M}(\lambda) = t\frac{r}{C}. \end{aligned} \tag{20}$$

$$\mathbb{D}^2(\xi_t) = \sum_{n=1}^{\infty} n^2 P_n(t) - t^2\frac{r^2}{C^2} = t\frac{r}{C}\left(1+\frac{t}{C}\right). \tag{21}$$

Let us now consider the secondary process generated by a mixed Poisson process. We suppose that every event is a starting point of a further happening the duration of which is also a random variable. If we measure this duration on lines going out from the points where the events occurred and drawn vertically to the time axis then we obtain a random point distribution on the plane. If λ is a constant then according to the results of the paper [3], this is a random point distribution of Poisson type. More precisely, if $\eta(A)$ denotes the number of points lying in the plane Borel-set A then to disjoint sets $A_1, \dots, A_r$ there correspond independent random variables $\eta(A_1), \dots, \eta(A_r)$ and $\eta(A)$ has a Poisson distribution with

$$\mathbb{M}(\eta(A)) = \lambda \iint_A d_x F(x,t)\,\mathrm{d}t, \tag{22}$$

provided (22) is finite, where $F(x,t)$ is the distribution function of the secondary happening starting at the time point t. If $F(x,t)$ has a density

$$f(x,t) = F'_x(x,t) \tag{23}$$

then

$$\mathbb{M}(\eta(A)) = \lambda \iint_A f(x,t)\,\mathrm{d}x\,\mathrm{d}t. \tag{24}$$

For the application of formulas (22), (24) see [3].

Now we can generalize this result to the case of a random λ. In this case clearly

$$\mathbb{P}(\eta(A_1) = k_1, \dots, \eta(A_r) = k_r \mid \lambda) = \prod_{i=1}^{r} \mathbb{P}(\eta(A_i) = k_i \mid \lambda) \tag{25}$$

provided $A_1, \dots, A_r$ are disjoint plane Borel-sets. Also we have

$$\mathbb{M}(\eta(A) \mid \lambda) = \lambda \iint_A d_x F(x,t)\,\mathrm{d}t, \tag{26}$$

(here we have supposed that the duration of the secondary happening do not depend on λ) hence $\eta(A)$ has a mixed Poisson distribution, i.e.

$$\mathbb{P}(\eta(A) = n) = \int_0^{\infty} \frac{(\lambda\nu^*(A))^n}{n!} e^{-\lambda\nu^*(A)}\,\mathrm{d}U(\lambda), \qquad n = 0,1,2,\dots, \tag{27}$$

where

$$\nu^*(A) = \iint_A d_x F(x,t)\,\mathrm{d}t. \tag{28}$$

If $U(\lambda)$ is the distribution function with

$$u(\lambda) = U'(\lambda) = C^r \frac{\lambda^{r-1}}{\Gamma(r)} e^{-C\lambda}, \qquad C > 0, \quad r > 0 \tag{29}$$

then

$$\mathbb{P}(\eta(A) = n) = \left(\frac{C}{C + \nu^*(A)}\right)^r \left(\frac{\nu^*(A)}{C + \nu^*(A)}\right)^n \frac{r(r+1)\cdots(r+n-1)}{n!} \qquad n = 0, 1, 2, \ldots \tag{30}$$

and

$$\mathbb{M}(\eta(A)) = \nu^*(A)\frac{r}{C}, \tag{31}$$

$$\mathbb{D}^2(\eta(A)) = \nu^*(A)\frac{r}{C}\left(1 + \frac{\nu^*(A)}{C}\right). \tag{32}$$

The underlying random process (the events of which are the starting points of the secondary happenings) is a Pólya-process if $C = r = \frac{1}{b}$, $b > 0$. In this case

$$\mathbb{P}(\eta(A) = n) = \left(\frac{\nu^*(A)}{1 + b\nu^*(A)}\right)^n \frac{(1+b)\cdots(1+(n-1)b)}{n!} \frac{1}{(1 + b\nu^*(A))^{\frac{1}{b}}},$$
$$n = 0, 1, 2, , \ldots . \tag{33}$$

With the aid of this model of secondary processes we can solve a number of special problems. Suppose e.g. that the underlying event process is the process of the calls arriving at a telephone centre and the secondary happenings are the conversations. Suppose moreover that the number of lines in the centre is infinite and the calls arrive according to a mixed Poisson process. Let furthermore $F(x,t) = 1 - e^{-\beta x}$, $\beta > 0$, $x > 0$. We want to determine the distribution of the random variable η_t denoting the number of conversations going on at time t. Clearly

$$\eta_t = \eta(A)$$

if $A = \{(\tau, x) : 0 \le \tau < t,\ x > t - \tau\}$ and thus

$$\mathbb{P}(\eta_t = n) = \mathbb{P}(\eta(A) = n) = \int_0^\infty \frac{(\lambda\nu^*(A))^n}{n!} e^{-\lambda\nu^*(A)}\, \mathrm{d}U(\lambda), \tag{34}$$

where

$$\nu^*(A) = \iint_A \beta e^{-\beta x}\, \mathrm{d}x\, \mathrm{d}t = \frac{1}{\beta}(1 - e^{-\beta t}). \tag{35}$$

It is not difficult to complete this results with the case when the process of calls is considered on the whole time axis $-\infty < t < \infty$. Then we interpret the probability (6) as the probability of arriving n calls in a time interval with the length t. If η_t denotes the same as before then $\eta_t = \eta(A)$ where $A = \{(\tau, x) : \tau < t,\ x > t - \tau\}$. Formula (26) remains true but $\nu^*(A)$ in this case equals $\frac{1}{\beta}$. If $U(\lambda)$ is given by (29) and $C = \frac{1}{b}$, $r = \frac{a}{b}$, $a > 0$, $b > 0$, then in view of (30) the probability that at time t there are n conversations, is

$$\mathbb{P}(\eta_t = n) = \frac{1}{(\beta + b)^n} \frac{(a+b)\cdots(a+(n-1)b)}{n!} \left(\frac{\beta}{\beta + b}\right)^{\frac{a}{b}}, \qquad n = 0, 1, 2, \ldots . \tag{36}$$

When $b \to 0$, this reduces to the Poisson distribution with the parameter $\dfrac{a}{\beta}$.

If $T = (-\infty, \infty)$, λ is a constant and $F(x,t)$ is independent of t then we know that η_t is a stationary process in the strict sense. This property remains true if λ is a random variable. In fact

$$\begin{aligned}\mathbb{P}(\eta_{t_1+\tau} = k_1, \ldots, \eta_{t_n+\tau} = k_n) &= \int_0^\infty \mathbb{P}(\eta_{t_1+\tau} = k_1, \ldots, \eta_{t_n+\tau} = k_n \mid \lambda)\,\mathrm{d}U(\lambda) \\ &= \int_0^\infty \mathbb{P}(\eta_{t_1} = k_1, \ldots, \eta_{t_n} = k_n \mid \lambda)\,\mathrm{d}U(\lambda) \\ &= \mathbb{P}(\eta_{t_1} = k_1, \ldots, \eta_{t_n} = k_n).\end{aligned} \tag{37}$$

Let m denote the expectation of λ,

$$m = \int_0^\infty \lambda\,\mathrm{d}U(\lambda). \tag{38}$$

Then the covariance of η_{s+t} and η_t is given by

$$R(t) = \mathbb{M}[(\eta(A_{s+t}) - m\nu^*(A_{s+t}))(\eta(A_s) - m\nu^*(A_s))], \tag{39}$$

where $A_t = \{(\tau, x),\ \tau < t,\ x > t - \tau\}$. Since for a fixed λ and disjoint sets $B_1, \ldots, B_r$ the random variables $\eta(B_1), \ldots, \eta(B_r)$ are independent, it follows that

$$R(t) = m\nu(A_s A_{s+t}). \tag{40}$$

Suppose that $F(x,t) = 1 - e^{-\beta x}$, $\beta > 0$, $x > 0$, then we get

$$R(t) = \frac{m}{\beta} e^{-\beta|t|}. \tag{41}$$

Thus if $\xi_t = \eta_t - \mathbb{M}(\eta_t) = \eta_t - m\nu^*(A_t)$, then the best linear least squares prediction of ξ_{s+t}, based on the variables ξ_t, $\tau \leq s$, is given by

$$\xi_{s+t} = \xi_s e^{-\beta|t|}. \tag{42}$$

References

[1] Lundberg, O. (1940). *On random processes and their application to stickness and accident statistics*, Uppsala.

[2] Prékopa, A. (1956). On stochastic set functions I., *Acta Math. Acad. Sci. Hung.*, **7**, 215–263.

[3] Prékopa, A. (1958). On secondary processes generated by a random point distribution of Poisson type, *Annales Univ. Sci. Budapest*, Sectio Math., **1**, 153–170.