ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
VIII (1966)

S. K. SRINIVASAN and K. S. S. IYER (Madras)

RANDOM PROCESSES ASSOCIATED WITH RANDOM POINTS ON A LINE

1. Introduction. Let us consider a random distribution of points on a line of finite extension t. We may regard t(1) as a general parameter so that t may stand for the time in any particular example. The random points occurring on the line may represent the events that occur in the interval $(0, t)$. To be specific we shall consider a stochastic process of the Poisson type with parameter $\lambda(t)$ so that the average number of events that occur in the time interval $(0, t)$ is given by

$$\Lambda(t) = \int_0^t \lambda(\tau)\, d\tau \tag{1.1}$$

while the probability that n events occur in the time interval $(0, t)$ is given by

$$\pi(n, t) = e^{-\Lambda(t)}[\Lambda(t)]^n/n!\ (^2). \tag{1.2}$$

We assume that every random event in the Poisson process triggers a signal whose amplitude is given by $f(t, \eta)$ where t is measured from the time the signal is produced and η may be a random parameter characterizing the signal. The quantity that is of interest is the sum of the amplitudes of the signals which, for a typical realization of events at $t_1, t_2, \ldots, t_n$, is given by

$$\xi(t) = \sum_i f(t-t_i, \eta_i). \tag{1.3}$$

The η_i's are assumed to be statistically independent random variables having the same distribution function. The distribution of $\xi(t)$ and its moments have been studied by Takacs ([6]) and Ramakrishnan ([3]).

(1) t may represent the point t as well as the interval $(0, t)$, the distinction being apparent from the context.

(2) Throughout the paper we shall use the symbol π to denote any probability frequency function, the difference between different probability frequency functions being apparent from the context.

In the present paper, we propose to consider the following generalization of the process mentioned above. We assume that each random point (or event) not necessarily of the Poisson type triggers off a random process $y(t)$ which originates from that point and progresses with respect to the same parameter. The random processes arising from different points do not interfere with each other. For a typical realization of events at $x_1, x_2, \ldots, x_n$ the variable that is of interest is given by

$$u^R(t) = \sum_{i=1}^{n} y^R(t-x_i)H(t-x_i), \tag{1.4}$$

where $y^R(t-x_i)$ is a typical realized value of the random variable $y(t-x_i)$ and $H(x)$ is heaviside unit function. At the point we observe that the probability measure of (1.4) involves detailed consideration of the probability measures of $y_1(t-x_1)$, $y_2(t-x_2), \ldots$ and $y_n(t-x_n)$. Thus $u(t)$ cannot represent a Markovian process and any direct attempt to obtain the Kolomogorov partial differential equation for the probability frequency function of $u(t)$ will prove futile. However, it is possible as will be shown in section 2 to obtain the moments of $u(t)$ by detailed consideration of the correlations of the random points.

If we confine our attention to 'renewal' processes (see for example Bartlett [1], p. 20), it is also possible to obtain integral equations for the probability frequency function of $u(t)$. This will be demonstrated in section 3 and it is worth mentioning that the integral equation is intractable and only some information regarding the moments may be extracted from the same. In the last section we shall deal with some applications to secondary Poisson events and theory of Carcinogenesis.

2. Moments of the distribution. The moments of the distribution of $u(t)$ can be studied with the help of certain correlation functions introduced in the theory of stochastic processes of continuous parametric systems. These correlation functions in some form or other have been studied some years ago in connection with the fluctuation problem of cosmic ray cascades and are known under different names such as cumulant functions, product densities, and Janossy densities (see for example Srinivasan [5]). Since our derivation of the moment formula depends on the product densities, we shall briefly recapitulate their definitions to suit our problem of random points on a line.

Let $N(x)$ represent the number of random points in the interval $(0, x)$. Then $dN(x)$ is the random variable representing the number of random points in the elemental range $(x, x+\delta x)$. We shall assume that the probability of finding one random point in $(x, x+\delta x)$ is proportional to δx so that the density relation

$$f_1(x)dx = E\{dN(x)\} \text{ [3]} \tag{2.1}$$

exists, the total probability of finding more than one random point in $(x, x+\delta x)$ being proportional to $o(\delta x)$. The integral of $f_1(x)$ over x yields only the mean number of random points in the range of integration, since the addition law of probability does not hold in its simple form, the events being in general not mutually exclusive. $f_1(x)$ is termed the product density of degree one. The product density of degree two is defined by

$$f_2(x_1, x_2)dx_1dx_2 = E\{dN(x_1)dN(x_2)\}, \qquad x_1 \neq x_2, \tag{2.2}$$

where $f_2dx_1dx_2$ represents the joint probability of finding two random points in the intervals (x_1, x_1+dx_1) and (x_2, x_2+dx_2) provided the two intervals do not overlap. However, when there is an overlap (i.e. $x_1 = x_2$)

$$E\{dN(x_1)dN(x_2)\} = E\{[dN(x_1)]^2\} = E\{dN(x_1)\} = f_1(x_1)dx_1; \tag{2.3}$$

such a simple result as (2.3) is due to the more general result that all the moments of the random variable $dN(x)$ are equal to the probability that the random variable takes the value 1. When similar degeneracies in x_i's occur in higher order product densities, we are led to lower order densities.

Let us assume that a point process is defined by the set of product densities as defined above. The random variable $u(t)$ can be expressed as

$$u(t) = \int_0^t dN(x, t)y(t-x)dx, \tag{2.4}$$

where t represents the 'duration' of the stochastic process (in the present case this being the size of the interval over which the coordinates x_i of the random points are distributed). The mean value of $u(t)$ can be readily obtained by taking the average of both sides of (2.4). Thus it follows from the linear character of the expectation operator

$$E\{u(t)\} = \int_0^t f_1(x, t)E\{y(t-x)\}dx. \tag{2.5}$$

The second moment of $u(t)$ is given by

$$E\{[u(t)]^2\} = \int_0^t\int_0^t f_2(x_1, x_2, t)E\{y(t-x_1)y(t-x_2)\}dx_1dx_2 + \int_0^t f_1(x, t)E\{[y(t-x)]^2\}dx. \tag{2.6}$$

---

[3] **Throughout this paper, E stands for expectation value of the quantity within the bracket.**

Since the $y(t-x_i)$ are statistically independent, $E\{y(t-x_1)y(t-x_2)\}$ can be written as $E\{y(t-x_1)\}E\{y(t-x_2)\}$, (2.6) being true even for the more general case when the variables $y(t-x_i)$ are correlated. To obtain higher moments, we have to classify the degeneracies arising from a product density of a given order (which is less than or equal to the order of the moment sought for). Suppose that in a product density of degree m, some intervals overlap leading to the product density of degree h. It is not sufficient to know the total number of ways in which such a degeneracy can arise. In fact, we should know the number of different complexions $C_h^m(i)$ leading to the product density of degree h, a complexion being characterized by a set of numbers $l_1, l_2, \ldots, l_h$ such that $l_1+l_2+ +\ldots+l_h = m$.

We note

$$\sum_i C_h^m(i) = C_h^m, \tag{2.7}$$

where the summation over i indicates the summation over different complexions. An explicit expression for $C_h^m(i)$ can be written down if we know many of the integers $l_1, l_2, \ldots, l_h$ are different. If only p out of the h integers $l_1, l_2, \ldots, l_h$ are different and $k_1, k_2, \ldots, k_p$ are the orders of the degeneracies in each of them, then

$$C_h^m(i) = \frac{m!}{l_1!\, l_2! \ldots l_h!\, k_1!\, k_2! \ldots k_p!}. \tag{2.8}$$

In fact, from combinatorial arguments it is easy to verify that $C_h^m(i)$ is the number of ways in which m distinguishable objects can be thrown into h groups of $l_1, l_2, \ldots, l_h$ objects respectively.

Thus the mth moment of $u(t)$ is given by

$$\begin{aligned} (2.9)\quad & E\{[u(t)]^m\} \\ &= \sum_{n=1}^{m} \sum_{l_1+l_2+\ldots+l_h=m} C_h^m(i) \int_0^t dx_1 \int_0^t dx_2 \ldots \int_0^t f_h(x_1, x_2, \ldots, x_h) \times \\ &\qquad \times E\{[y(t-x_1)]^{l_1}\} E\{[y(t-x_2)]^{l_2}\} \ldots E\{[y(t-x_h)]^{l_h}\} dx_h. \end{aligned}$$

If, however, the random variables $y(t-x_i)$ are statistically dependent, (2.9) still holds good provided we replace

$$E\{[y(t-x_1)]^{l_1}\} E\{[y(t-x_2)]^{l_2}\} \ldots E\{[y(t-x_h)]^{l_h}\}$$

by

$$E\{[y(t-x_1)]^{l_1} [y(t-x_2)]^{l_2} [y(t-x_h)]^{l_h}\}.$$

Thus the moments of $u(t)$ can be calculated if we know the distribution of the random points on the line or equivalently if we are given the product densities of the random points of different orders. We finally observe that there may be situations where y may depend also on the origin of the primary process. In this case, the formula (2.9) is equally true if we replace $y(t-x_i)$ by $y(x_i, t-x_i)$.

As an example, let us consider the case when the x_i's are a process with a stationary positive increments. More precisely we assume that the distribution of the time-interval between successive events in a simple renewal process is given by

$$\varphi(t) = 4\lambda^2 te^{-2\lambda t}. \tag{2.10}$$

The product densities of degree one and two can be found to be (see Bartlett [1], page 167)

$$f_1(t) = \lambda - \lambda e^{-4\lambda t}, \tag{2.11}$$

$$f_2(t_1, t_2) = \lambda^2 - \lambda^2 e^{-4\lambda(t_1-t_2)}, \quad t_1 > t_2. \tag{2.12}$$

If we further assume that each event at x_i triggers off a random process $y(t)$ which originates from that instant x_i of time, the random variable $y(t)$ assuming a value $e^{-n(t)}$ where $n(t)$ is a simple Poisson process progressing with respect to t with a parameter μ. The mean and mean square value of $u(t)$ are given by

$$E\{u(t)\} = \frac{\lambda}{\mu}(1-e^{-\mu t}) - \frac{\lambda}{\mu-4\lambda}(e^{-4\lambda t}-e^{-\mu t}), \tag{2.13}$$

$$E\{[u(t)]^2\} = \lambda^2\left[\frac{(1-e^{-\mu t})^2}{\mu^2} - \frac{1}{\mu^2-16\lambda^2}\{1-e^{-(\mu-4\lambda)t}-e^{-(\mu+4\lambda)t}-e^{-2\mu t}\}\right]. \tag{2.14}$$

It is interesting to note that as t becomes large, mean and mean square value of $u(t)$ tend to finite limits.

3. Probability frequency function of $u(t)$. As has been mentioned in section 1, $u(t)$ does not represent a Markovian process and as such it is not possible to write down the partial differential equation satisfied by $\pi(y, t)$. Apart from the nature of $y(t)$, the distribution of random points on the t-axis need not necessarily be Markovian. If $\pi(n, t)$ is the probability that n points have occurred between 0 and t the probability of occurrence of further random points between t and $t+dt$ depends not only on the number n but on the positions of the events on the t-axis as well. However, if we confine ourselves to a renewal process progressing with respect to t, it is possible to write down integral equations for the probability frequency function of $u(t)$.

Let $\varphi(t)$ be the probability that no random point occurs up to t given that a random point occurred at $t = 0$ and let $P(u, t)du$ be the probability that the random variable takes a value between u and $u+du$

given that a random point occurred at $t = 0$. For reasons that will be apparent presently, we shall, in dealing with $P(u, t)$, remove from $u(t)$ the random variable $y(t)$ generated by the random point at $t = 0$. To obtain $P(u, t)$ we need compound the probability frequency functions of $y(t)$ and $u(t)$.

To obtain an integral equation for $P(u, t)$ we focus our attention on the first random point that occurs after $t = 0$. The probability that the first random point occurs between τ and $\tau + d\tau$ is given by $-\varphi'(\tau)d\tau$. Thus $P(u, t)$ satisfies the equation

$$P(u, t) = -\int_y \int_0^t d\tau P(u-y, t-\tau)\varphi'(\tau)\pi(y, t-\tau)dy. \tag{3.1}$$

It is difficult to obtain explicit solution for $P(u, t)$ even in simple cases. Defining the Laplace transform of $P(u, t)$ as $P(\xi, t)$ we can write (3.1) as

$$P(\xi, t) = -\int_0^t P(\xi, t-\tau)\pi(\xi, t-\tau)\varphi'(\tau)d\tau. \tag{3.2}$$

The moments of u can be obtained from (3.2) by differentiating equation (3.2) with respect to ξ at the point $\xi = 0$ and solving the resulting equation recursively. The first three moments obtained from (3.2) are found be in agreement with formula (2.8).

4. Physical applications. As an example let us consider the secondary processes generated by a Poisson process. We shall assume that every Poisson event gives rise to a discrete Markovian process progressing with respect to the same parameter, the random variable $y(t)$ assuming only integral values. The symbol $n(t)$ will be used instead of $y(t)$ for obvious reasons and $\psi(n, t)$ will be assumed to be the probability frequency function of $n(t)$ so that the probability that $n(t) = n$ is just $\psi(n, t)$.

In this case it is easy to compute the probability that the random variable $u(t)$ assumes the integral value m by considering the joint distribution of the position of events that occur on the t-axis. The probability that the first event happens between t_1 and $t_1 + dt_1$, the second between t_2 and $t_2 + dt_2$, ... the ith event between t_i and $t_i + dt_i$ and no further event occurs up to t is given by

$$e^{-\lambda t_1}\lambda dt_1 e^{-\lambda(t_2-t_1)}\lambda dt_2 \ldots e^{-\lambda(t_i-t_{i-1})}\lambda dt_i e^{-\lambda(t-t_i)} = e^{-\lambda t}\lambda^i dt_1 dt_2 \ldots dt_i. \tag{4.1}$$

Thus $\pi(m, t)$ is given by

$$\pi(m, t) \tag{4.2}$$
$$= \sum_{\substack{n_1 n_2 \ldots n_i \\ n_1+n_2+\ldots+n_i=m}} e^{-\lambda t}\int_0^t dt_1 \int_{t_1}^t dt_2 \ldots \int_{t_{i-1}}^t dt_i \psi(n_1, t-t_1)\psi(n_2, t-t_2)\ldots\psi(n_i, t-t_i)\lambda^i.$$

Defining the generating function of $\pi(n, t)$ and $\psi(n, t)$ as $g(z, t)$ and $\chi(z, t)$, we obtain

(4.3) $g(z, t)$

$$= e^{-\lambda t} \sum_{n=0}^{\infty} \int_0^t dt_1 \int_{t_1}^t dt_2 \dots \int_{t_{n-1}}^t dt_n \lambda^n \chi(z, t-t_1)\chi(z, t-t_2)\dots\chi(z, t-t_n).$$

If we still further specialize and take $\chi(z, t)$ to be Poissonian with parameter μ we find

$$g(z, t) = e^{-\lambda t} \sum_{n=0}^{\infty} \int_0^t dt_1 \int_{t_1}^t dt_2 \dots \int_{t_{n-1}}^t dt_n \lambda^n \exp\left[-\mu(1-z) \sum_{i=1}^{n}(t-t_i)\right]. \tag{4.4}$$

The moments of n can be obtained from (4.4). Alternatively the moments can be obtained by making use of (2.5).

As a second example, let us consider a problem in Carcinogenesis. In cancer research, the observations are usually confined to a guineau pig (usually a mouse) on whom a certain amount of Carcinogen is administered (see for example Tucker [7]). The Carcinogen produces in the body of a mouse hyperplastic foci according to certain probability laws. Each hyperplastic focus passes through two stages in its mode of development and becomes a tumour. And hyperplastic focus has a certain probability of anihilation in either stage of its life and the conversion of a hyperplastic focus into a tumour is only a competitive process in the second stage of its life. However, once a tumour is formed, it can never be removed and the biological interest centers round the growth of these tumours and the size of each tumour taking into account the number of cells of which the tumour is composed. Since the administration of carcinogen produces only local effects, the random variable that is of interest is the total number of cells of which the tumours are composed. The times of formation of tumours correspond to the random points of section 1 and the growth of the cells in a tumour correspond to the random process associated with the random point. Thus the formula (2.6) may be used to obtain the moments of the number of cells formed. In the model used by Tucker, it was assumed that the cells in any individual tumour increase in a Poissonian way with a time dependent parameter, the time being measured from the time of administration of Carcinogen. However, a more realistic model will be one in which the growth depends also on the time of formation of the tumour. Such a model exactly fits in with the stochastic process outlined in section 2. Thus the formula (2.8) can be used for the calculation of the moments of the total number of cells, of course the determination of the product densities of various orders of the times of formation of tumours is an independent problem falling outside the scope of the present work.

Finally, we shall outline a simple problem in the statistical theory of brightness of the Milky Way. Some years ago, Chandrasekhar and Münch ([2]) examined some interesting problems of Stellar Statistics arising from the distribution of inter-stellar matter. It was assumed that inter-stellar matter occurs in the form of discrete clouds, the clouds being characterized by a distribution function governing the optical thickness of the clouds. The contribution to the intensity of brightness arising from the stars was assumed to be uniform and proportional to the extent of the astrophysical system. In a more realistic model, the discrete nature of the stars must be incorporated. In fact, there are two aspects of the problem. The first relates to the distribution of the stars along the line of observation while the second, to the distribution of intensities of the individual stars. We shall assume that the distribution of stars along the line of sight (t-axis) is specified by the various product densities $f_n(x_1, x_2, \ldots, x_n)$ $(n = 1, 2, \ldots)$ and that the intensities are independently distributed, the probability frequency function governing the intensity I of any star being given by $h(I)$. In the notation of section 1, each random point x corresponds to the cumulative transparency factor Q corresponding to the extent x of the astrophysical system since the intensity I of the star occurring at x is only observed to be IQ at t. Thus $u(t)$ represents the total intensity measured by an observer stationed at the origin. However, (2.9) cannot be used for the moments of $u(t)$ since the transparency factors $Q_1, Q_2, \ldots, Q_n$ do not represent independent random processes. Nevertheless, the calculations up to (2.9) can be taken as applicable in the present case except that the product

$$E\{[y(t-x_1)]^{l_1}\}E\{[y(t-x_2)]^{l_2}\}\ldots E\{[y(t-x_i)]^{l_i}\}$$

is to be replaced by

$$\int_{Q_1}\int_{Q_2}\ldots\int_{Q_i} E\{I^{l_1}\}E\{I^{l_2}\}\ldots E\{I^{l_i}\}\times$$
$$\times Q_1^{l_1}Q_2^{l_2}\ldots Q_i^{l_i}\pi(Q_1, Q_2, \ldots, Q_i, x_1, x_2, \ldots, x_i)\,dQ_1\,dQ_2\ldots dQ_i.$$

The integrations over $Q_1, Q_2, \ldots, Q_i$ can be performed and we shall not explicitly calculate them since the method of calculation of such integrals is straightforward (see for example Ramakrishnan and Srinivasan [4]).

In the last example, incidentally we have also indicated how we can deal with certain situations where the secondary random processes are correlated with each other. In fact, the moment formula (2.8) can be generalized with obvious modifications to include the correlation between different y's.

References

[1] M. S. Bartlett, *An introduction to stochastic processes*, Cambridge University Press (1955).

[2] S. Chandrasekhar and G. Munch, *Theory of fluctuations in the brightness of Milky Way*, Ap. J. 112 (1950), pp. 380-392.

[3] A. Ramakrishnan, *Stochastic processes associated with random divisions of a line*, Proc. Camb. Phil. Soc. 49 (1953), pp. 473-485.

[4] A. Ramakrishnan and S. K. Srinivasan, *Correlation problems in the brightness of Milky Way*, Ap. J. 123 (1956), pp. 479-485.

[5] S. K. Srinivasan, *Multiple stochastic point processes*, Zastosow. Matem. 6 (1962), pp. 209-219.

[6] L. Takacs, *On secondary processes generated by a Poisson process and their applications in Physics*, Act. Math. Acd. Sci. Hung. 5 (1954), pp. 203-236.

[7] H. J. Tucker, *A stochastic model for a two stage theory of carcinogenesis*, Proc. Fourth. Berkeley. Symp. Math. Statist. 4 (1961), pp. 387-403.

INDIAN INSTITUTE OF TECHNOLOGY, DEPARTMENT OF MATHEMATICS

Received on 18. 3. 1965

S. K. SRINIVASAN i K. S. S. IYER (Madras)

PROCESY STOCHASTYCZNE ZWIĄZANE Z LOSOWYMI PUNKTAMI NA PROSTEJ

STRESZCZENIE

Procesy stochastyczne związane z losowymi punktami na prostej odgrywają wielką rolę w zagadnieniach przyrodniczych. W niniejszej pracy autorzy zajmują się procesami związanymi z losowymi punktami na skończonym odcinku osi t. Przyjmuje się, że każdy z punktów losowych, które to punkty nie muszą być rozłożone według rozkładu Poissona, wyzwala proces $y(t)$. Wyzwolony proces począwszy od danego punktu rozwija się losowo wraz z parametrem t, przy czym procesy wyzwolone w różnych punktach mogą być lub nie być wzajemnie niezależne. Proces wyzwolony w punkcie t_i generuje sygnał $y(t-t_i)$ określony dla chwil późniejszych. Przedmiotem pracy jest badanie statystycznych własności sumy $u(t)$ wszystkich sygnałów generowanych w chwili t przez procesy wyzwolone przed tą chwilą. Proces $u(t)$ nie jest na ogół procesem Markowa i próby bezpośredniego wyznaczenia dystrybuanty tego procesu są bezowocne. Można jednak wyznaczyć wszystkie momenty procesu $u(t)$ stosując metody badania procesów punktowych i próbując wyrazić korelację sumy sygnałów w zależności od korelacji poszczególnych sygnałów jak również korelacji punktów losowych na osi czasowej. Ponadto, dla procesów odnowy można wyprowadzić równania całkowe na gęstość rozkładu prawdopodobieństwa.

Autorzy ilustrują proponowaną metodę formułując kilka zagadnień z zakresu wtórnych procesów Poissona. W zakończeniu pracy autorzy omawiają krótko zastosowania do innych zjawisk przyrodniczych, takich jak korelacje intensywności świecenia Drogi Mlecznej i rozkłady ilości komórek rakowych w procesie karcynogenezy.

С. К. СРИНИВАСАН и К. С. С. ИЕР (Мадрас)

СТОХАСТИЧЕСКИЕ ПРОЦЕССЫ СВЯЗАННЫЕ СО СЛУЧАЙНЫМИ ТОЧКАМИ НА ПРЯМОЙ

РЕЗЮМЕ

Стохастические процессы связанные со случайными точками на прямой играют значительную роль в исследованиях в естествознании. В этой статье авторы рассматривают модель стохастических процессов связанных со случайными точками на конечном интервале оси t. Предполагается что каждая случайная точка вызывает случайный процесс $y(t)$. Начиная с этой точки процесс развивается с течением параметра t, но отдельные процессы вызваны разными случайными точками могут быть, или не быть взаимно независимыми. Процесс вызванный в точке t_i порождает сигнал $y(t-t_i)$ определенный для всех последующих моментов t за точкой вызова. В статье авторы исследуют статистические свойства суммы $u(t)$ сигналов порожденных в моменте t всеми процессами, вызванными до этого момента. Процесс $u(t)$ как правило не является процессом Маркова и попытки непосредственного определения его функции распределения оказываются безполезными. Но удается определить все моменты процесса $u(t)$ применяя методы исследования точечных процессов и пытаясь установить зависимость корреляции суммы от корреляций отдельных сигналов и корреляции случайных точек на оси. Для процессов восстановления можно также вывести интегральные уравнения на плотность распределения вероятностей.

Авторы дают примеры формулировки некоторых задач из области вторичных процессов Пуассона. В заключении приведено несколько применений метода к другим проблемам естествознания: к исследованию корреляции интенсивности света Молочной Дороги и распределения числа раковых клеток в процессе канциногенезы.