ON A ZERO-ONE PROCESS AND SOME OF ITS APPLICATIONS

By HIROTUGU AKAIKE

(Received Nov. 15, 1956)

0. Introduction and summary

We have many occasions to analyse the pattern of occurrences of serial random events, for example, automobile flows, dropping ends of cocoon filaments. Many works have been done for the case when the lengthes of gaps independently follow one and the same negative exponential distribution. In this case, as is well known, the number of occurrences in one time interval is entirely independent of that in another time interval which does not overlap the former, and follows the Poisson distribution. Nevertheless, we sometimes observe random events concerning which the numbers of occurrences in some disjoint time intervals are not entirely independent, that is, they show wave-like movement. In some cases, this is due to the fact that the distribution of gaps are not necessarily of negative exponential type. In this paper we shall treat a discrete parameter processes for the purpose to analyse practically the processes of the type just stated.

The process we treat here will be a powerfull tool for analysis of the fundamental strucutre of the process representing the dropping ends of cocoon filaments. Some examples of its application will also be given.

1. Definition of gap-process

Consider a stochastic process $\{x_n(\omega)\,;\ n=0,\ 1,\ 2,\ \cdots\}$, where $x_0(\omega)\equiv 1$ and $x_n(\omega)=1$ or 0, and which is governed by the gap distribution $\text{Prob}\,\{x_{n+1}=0,\ x_{n+2}=0,\ \cdots x_{n+\nu-1}=0,\ x_{n+\nu}=1|x_n=1\}\equiv p_\nu\ (\nu=1,\ 2,\ \cdots)$.

Then we have $P_n\equiv\text{Prob}\,\{x_n(\omega)=1\}=\sum_{\nu=1}^{n} p_\nu\, P_{n-\nu}$ where $P_{n-\nu}$ is given by the recurrence relation

$$P_0\equiv 1, \qquad P_\mu=\sum_{\nu=1}^{\mu} p_\nu\, P_{\mu-\nu}\,.$$

As for the joint distribution of $x_n(\omega)$'s, from the relation

$$\text{Prob}\{x_{n+\nu}(\omega)=1|x_n(\omega)=1,\ x_{n-1}(\omega),\ \cdots,\ x_0(\omega)\}$$
$$=\text{Prob}\{x_{n+\nu}(\omega)=1|x_n(\omega)=1\}=P_\nu$$

we have

$$\text{Prob}\{x_n(\omega)=1,\ x_{n+\nu_1}(\omega)=1,\ x_{n+\nu_1+\nu_2}(\omega)=1,\ \cdots\ x_{n+\nu_1+\nu_2+\cdots+\nu_k}(\omega)=1\}$$
$$=P_n P_{\nu_1} P_{\nu_2} \cdots P_{\nu_k}{}^*.$$

Now, as is well known, by the recurrent event theory, when the greatest common divisor of those ν for which $p_\nu > 0$ is one, we have

$$\lim_{n\to\infty} P_n = P = \frac{1}{L} \quad \text{where} \quad L=\sum \nu p_\nu$$

(see [1]). Hence, when we are only concerned with the steady state of this process after a long time has passed the process may be represented by another strictly stationary zero-one process which is defined as follows.

We call the stochastic process $\{x_n(\omega),\ -\infty < n < \infty\}$ a gap process where for almost all ω $x_n(\omega)=0$ or 1 and

$$\text{Prob}\{x_n(\omega)=1,\ x_{n+\nu_1}(\omega)=1,\ \cdots,\ x_{n+\nu_1+\cdots+\nu_k}(\omega)=1\}=PP_{\nu_1}\cdots P_{\nu_k}$$

where P_ν is given from some $\{p_\nu\}$ with $\sum_{\nu\geq 1} p_\nu=1$ and $p_\nu \geq 0$ by the recurrence relation $P_0\equiv 1$, $P_\nu=\sum_{\mu=1}^{\nu} p_\mu P_{\nu-\mu}$ and $P=\lim_{\nu\to\infty} P_\nu$. We shall call $\{p_\nu\}$ the corresponding gap distribution. Its meaning will be obvious from the former consideration.

2. Spectral function of bounded gap-process

In this section we shall study some of the spectral character of the gap process whose gap lengthes are bounded by some fixed value. The restriction of boundedness for the gap lengthes is not a serious one for practical problems. Now, let the gap distribution be $\{p_\nu;\ \nu=1,2,\cdots,k.\ \sum_{\nu=1}^{k} p_\nu=1\}$. Namely, consider the gaps not longer than k. Then we have the following:

THEOREM. *When* $p_{\nu_i}\neq 0\ (i=1,\cdots,\mu)$ *and* $\sum_{i=1}^{\mu} p_{\nu_i}=1$ *hold and the*

---

* To determine the joint distribution of zero-one variables only the values of the type $\text{Prob}\{x_{n_1}(\omega)=1,\ x_{n_2}(\omega)=1\ \cdots\ x_{n_k}(\omega)=1\}$ are needed.

For example, to get the value of $\text{Prob}\{x_{n_1}(\omega)=1,\ x_{n_2}(\omega)=0,\ x_{n_3}(\omega)=1\}$ the following relation can be used.

$$\text{Prob}\{x_{n_1}(\omega)=1,\ x_{n_2}(\omega)=0,\ x_{n_3}(\omega)=1\}$$
$$=\text{Prob}\{x_{n_1}(\omega)=1,\ x_{n_2}(\omega)=1\}-\text{Prob}\{x_{n_1}(\omega)=1,\ x_{n_2}(\omega)=1,\ x_{n_3}(\omega)=1\}.$$

greatest common divisor of $(\nu_1, \nu_2, \nu_3, \cdots, \nu_\mu)$ *is* 1, *the* $\{x_n(\omega)-P\}$ *process has the continuous spectral density function*

$$G'(\lambda)=2R(0)+4\sum_{n=1}^{\infty}R(n)\cos 2\pi n\lambda$$

where
$$R(n)=P(P_n-P)\,.$$

PROOF. As is easily seen, the covariance function of the $\{x_n(\omega)-P\}$ process is given by $R(n)=E\{x_{n+m}(\omega)-P\}\{x_m(\omega)-P\}=P(P_n-P)$. Thus the process is stationary in the wide sense. Now, as the P_n's are the solution for the difference equation.

$$P_n=p_1P_{n-1}+p_2P_{n-2}+\cdots+p_kP_{n-k}$$

under the initial condition

$$P_0=1,\ P_1=p_1,\ P_2=p_1P_1+p_2,\ \cdots,\ P_{k-1}=p_1P_{k-2}+\cdots+p_{k-1},$$

it can be represented in the form

$$P_n=c_1z_1^n+c_2z_2^n+\cdots+c_kz_k^n$$

where z_i's are the roots of characteristic equation $z^k-p_1z^{k-1}-p_2z^{k-2}-\cdots-p_k=0$ and c_i's are complex numbers. We can take $z_1=1$, as 1 is clearly a root of the characteristic equation. Moreover, it is shown from the following inequality that 1 is a simple root of the characteristic equation:

$$\frac{d}{dz}(z^k-p_1z^{k-1}-p_2z^{k-2}-\cdots-p_k)\Big|_{z=1}$$
$$=k\Big(1-\frac{k-1}{k}p_1-\frac{k-2}{k}p_2-\cdots-\frac{1}{k}p_{n-1}-\frac{0}{k}p_k\Big)=\sum_{\nu=1}^{k}\nu p_\nu>0\,.$$

If $|z_j|=1$ holds for some j, we can represent it by $z_i=e^{i\theta}$, and this θ must satisfy the following equation

$$1=p_1e^{-i\theta}+p_2e^{-i2\theta}+\cdots+p_ke^{-ik\theta}$$
$$=p_{\nu_1}e^{-i\nu_1\theta}+p_{\nu_2}e^{-i\nu_2\theta}+\cdots+p_{\nu_\mu}e^{-i\nu_\mu\theta}\,.$$

Hence we must have the following equations

$$e^{-i\nu_j\theta}=1 \qquad j=1,2,\cdots,\mu\,,$$

that is,

$$\nu_j\theta=\kappa_j2\pi \qquad (\kappa_j \text{ integer})\,.$$

Now, as the greatest common divisor of $\nu_1, \nu_2, \cdots, \nu_\mu$ is unity, we have for some integral values $\alpha_1, \alpha_2, \cdots, \alpha_\mu$

$$1=\alpha_1\nu_1+\alpha_2\nu_2+\cdots+\alpha_\mu\nu_\mu .$$

Thus we have

$$\theta=(\alpha_1\kappa_1+\alpha_2\kappa_2+\cdots+\alpha_\mu\kappa_\mu)2\pi$$

or $z_i=1$. From these facts we get a conclusion that $|z_i|<1$ for $i\neq 1$. Under the condition of the theorem we must have

$$\lim_{n\to\infty} P_n=P=\frac{1}{L}$$

where $L=\sum \nu p_\nu$, and from this it can be seen that $c_1=P$ and

$$P_n-P=\sum_{i=2}^{k} c_i z_i^n \qquad |z_i|<1 .$$

Therefore, we have

$$\sum_{n=0}^{\infty}|R(n)|=P\sum_{n=0}^{\infty}|P_n-P|\leq P\sum_{n=0}^{\infty}\sum_{i=2}^{k}|c_i||z_i|^n<+\infty$$

and the existence of the continuous spectral density function

$$G'(\lambda)=2R(0)+4\sum_{1}^{\infty}R(n)\cos 2\pi n\lambda$$

is assured [2].

From this theorem it follows that almost all gap-processes we observe have continuous spectral functions and that some of them show wave-like movements. In such cases, according to the theory of stationary process, the ordinary periodogram analysis is of less use, and the structure of gap distribution must be considered first. When $R(n)$'s $(n=0, 1\cdots, k)$ are given, we can easily obtain the values $\{p_\nu ; \nu=1, 2, \cdots k\}$ by using the recurrence relation $p_\nu=P_\nu-\sum_{\mu=1}^{\nu-1}p_\mu P_{\nu-\mu}$.

3. Examples. (Practical application of gap process theory).

The examples we treat in this section are essentially continuous parameter processes. But as the treatment of such processes are not easy we have applied discrete approximations to them which we believe to have sufficient accuracy for practical purposes. Original processes are continuous parameter processes of which almost all sample functions are step-functions with jumps of unit saltus, and the time lengthes τ between jumps are governed independently of each other by one and the same continuous probability distribution $p(\tau)d\tau$.

Let τ_0 and ε be positive numbers such that

$$\int_0^{\tau_0} p(t)dt \leq \varepsilon \qquad \frac{T}{\tau_0}\varepsilon \ll 1$$

where T is the whole length of the time for observation. When we put $\varepsilon(\tau) \equiv \int_0^{\tau} p(t)dt$, we have for continuous $p(t)$ a positive h such that

$$\varepsilon(\tau) = \tau p(h\tau) \qquad 0 < h < 1 ,$$

so, whenever $\lim_{t \to 0} p(t) = 0$, we can find a pair of ε and τ_0 which satisfies the above stated conditions. This means that for continuous $p(\tau)$ with $\lim_{t \to 0} p(t) = 0$, there exists such τ_0 that there are few intervals containing more than two occurrences of events in some finite number of observations each with the duration of time T. To treat such processes the following procedure will always be a pertinent one.

Given a record of occurrences of events, we take as the time unit for digitization such τ' that there is no time interval of length τ' containing more than two records of occurrences.

1. *Flow of automobiles.*

We have observed the flow of automobiles at some points in Tokyo. The distribution of the number of cars per unit time interval seems to be fairly well fitted by the Poisson distribution (Fig. 1). Nevertheless, when we analyse the dependence between the numbers of cars in succesive time intervals we can recognize the existence of serial correlations. This fact contradicts to the fundamental assumption that the numbers of cars in mutually disjoint time intervals are independent of each other for a simple Poisson process. Therefore, we have digitized the data by taking τ' as half the shortest record of time interval between the cars, and analysed the distribution of the length of time intervals between

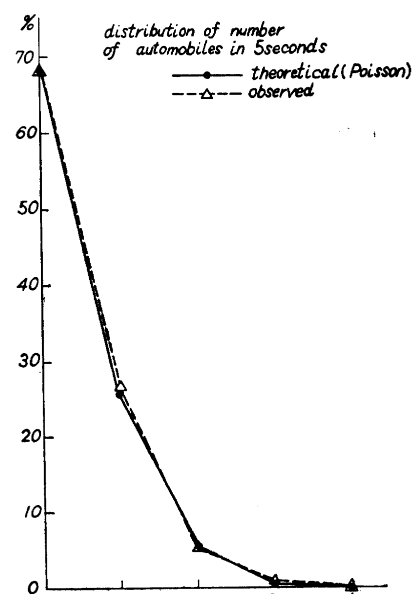

Fig. 1.

successive cars and got the result as shown in Fig. 2.

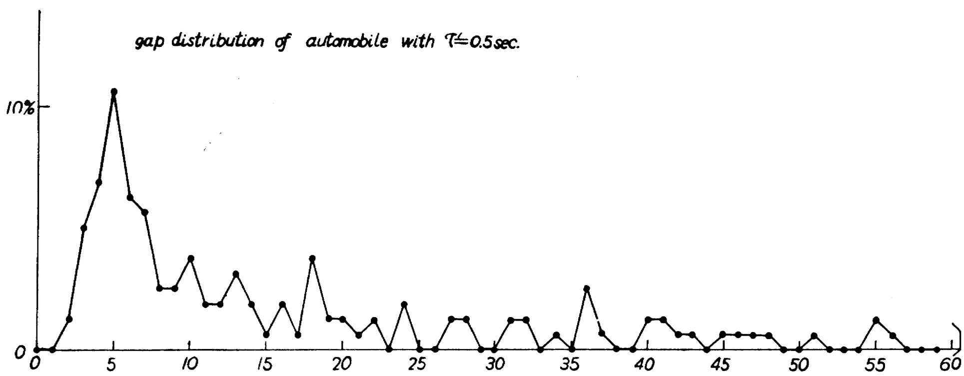

Fig. 2.

This figure suggests that the above hypothesis does not hold true, that is, the length of time interval between successive cars does not follow the negative exponential distribution. We have taken the Fig. 2 as giving the gap distribution $\{p_\nu\}$ and calculated P_ν's under the hypothesis that the process is a gap one. The result is graphically shown in Fig. 3. This shows a fairly good fit and the existence of serial correlation between successive data can be considered as due to the narrowness of the road. We observed 156 cars in this experiment.

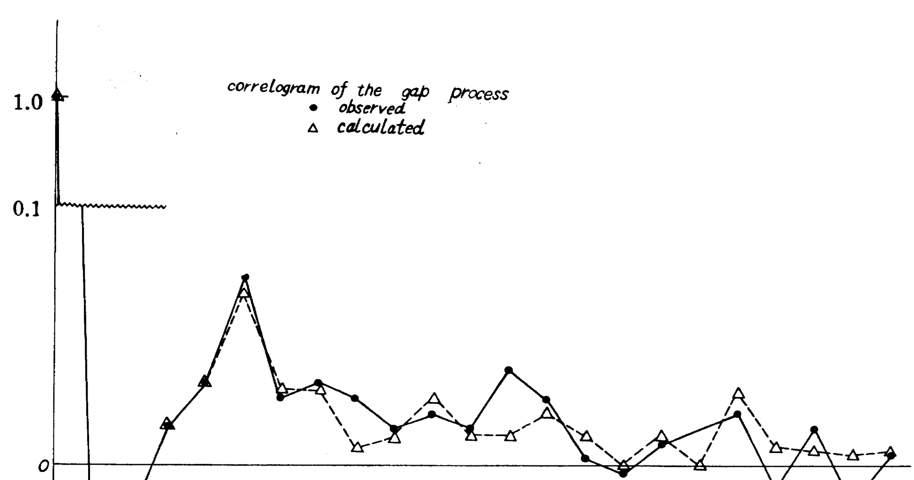

Fig. 3.

2. *Dropping ends of cocoon filament.*

When we operate the silk-reel we sometimes observe the wave-like pattern of stoppings by dropping ends of cocoon. This seems due to the fact that the distribution of length between successive ends does not follow the negative exponential distribution. Experimental data show that the distribution of nonbreaking length of a bave is such as shown in Fig. 4*.

Taking this figure as giving a gap distribution $\{p_\nu\}$ we get the sequence $\{P_\nu\}$. The values $(P_\nu - P)/(1-P)$ are considered as showing the correlogram of the process** (Fig. 5).

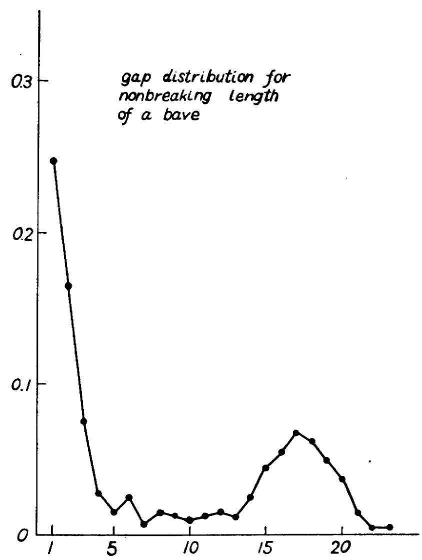

Fig. 4.

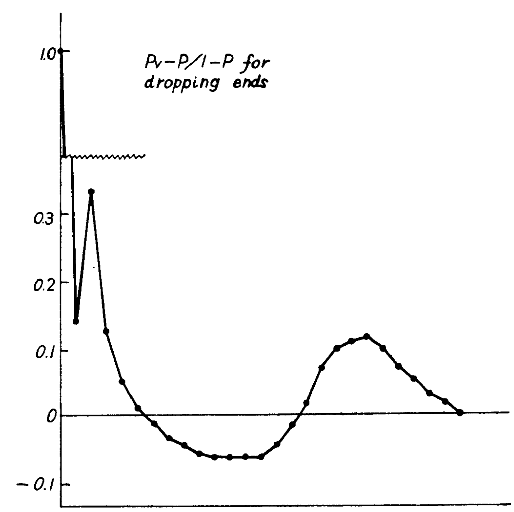

Fig. 5.

4. Acknowledgement

The author expresses his thanks to Mr. K. Morita, member of Electrical Communication Laboratory, Nippon Telegraph and Telephone Public Corporation, who kindly provided us with the convenience to use a recording apparatus for our experiments on automobile flows and to Mr. A. Shimazaki* for calling the author's attention to the problem of dropping ends of cocoon filaments.

* The data is offered to the author from Mr. A. Shimazaki of Sericultural Experiment Station, Tokyo.

** This line of approach seems to give some light on the development of the production of raw silk, and is now being put forward by Mr. A. Shimazaki.

Thanks are also due to Mr. M. Sibuya of our Institute for kind discussion on the dropping ends problem and to Misses Y. Saigusa and T. Kageyama for their recording and arranging work of data.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] W. Feller, *An Introduction to Probability Theory and its Applications*. John Wiley & Sons, Inc., New York, (1950), pp. 243-248.

[2] J. L. Doob, *Stochastic Processes*. John Wiley & Sons, Inc., New York, (1953), p. 476.