ON RANDOM SAMPLING FROM A STOCHASTIC PROCESS[1]

By J. R. Blum and Judah Rosenblatt

University of New Mexico and Sandia Corporation

1. The problem. Let $\{X_n, n = 1, 2, \cdots\}$ be a stochastic process which is stationary and ergodic. Then it follows from the individual ergodic theorem that we may estimate the entire probability structure of the process by an observation $\{x_1, x_2, \cdots\}$ on the process.

Assume from now on that the random variables of the process are two-valued, i.e. $P\{X_n = 0\} = 1 - p = 1 - P\{X_n = 1\}$, where $0 < p < 1$. This is an unessential restriction which serves to simplify the ideas involved.

Now suppose that there are physical reasons which prohibit us from observing each of the successive random variables X_n. If we are then forced to observe a subsequence $\{X_{k_n}, n = 1, 2, \cdots\}$, we may ask whether it is still possible to estimate the probability structure of the original process from observing $\{X_{k_1}, X_{k_2}, \cdots\}$. In general the answer to this question is in the negative. For example, if k is an integer, $k > 1$, and $k_n = kn$ for $n = 1, 2, \cdots$; then while the process $\{X_{k_n}\}$ is still stationary and may also be still ergodic, it may be impossible to estimate the joint distribution of X_1 and X_2. Moreover in general the process $\{X_{k_n}\}$ may not be ergodic, or even stationary.

In this paper we shall consider what can be done with random sampling, that is when we assume that $\{k_n\}$ is a sequence of random variables. To formalize this notion we shall assume that in addition to the $\{X_n\}$ process we have at our disposal a sequence of random variables $\{Y_n, n = 1, 2, \cdots\}$ where the $\{Y_n\}$ process is independent of the $\{X_n\}$ process and consists of positive integer-valued random variables. We shall assume throughout that the $\{Y_n\}$ process is a stationary, ergodic process. In terms of the bivariate $\{X_n, Y_n\}$ process we can define a new bivariate process $\{Y_n, Z_n\}$ where the $\{Y_n\}$ process is as above and $Z_n = X_{N(n)}$, where $N(n) = \sum_{j=1}^{n} Y_j$, $n = 1, 2, \cdots$, and we assume that it is the $\{Y_n, Z_n\}$ process which is being observed. As we shall see below, it is easy to prove that the $\{Y_n, Z_n\}$ process is stationary. Under certain assumptions it will be shown that this process is also ergodic. Assume for the moment that this has already been done. Define

$$\begin{aligned} f(y, z_1, z_2) &= 0, && \text{if } y \neq 1 \\ &= z_1 z_2, && \text{if } y = 1. \end{aligned}$$

Then

Received January 9, 1964.

[1] Research supported by the National Science Foundation, Grant GP-1816, and in part by the United States Atomic Energy Commission.

$$Ef[Y_2, Z_1, Z_2] = P\{Y_2 = 1\}E\{Z_1Z_2 \mid Y_2 = 1\}$$
$$= P\{Y_2 = 1\}\sum_{i=1}^{\infty} P\{Y_1 = i \mid Y_2 = 1\}E\{X_iX_{i+1} \mid Y_1 = i, Y_2 = 1\}$$
$$= P\{Y_2 = 1\}P\{X_1 = 1, X_2 = 1\}.$$

Now let

$$\begin{aligned} \epsilon_j &= 1, \quad &&\text{if } Y_j = 1 \\ &= 0, &&\text{otherwise,} \end{aligned}$$

and define W_n by

$$W_n = \sum_{j=1}^{n} f[Y_{j+1}, Z_j, Z_{j+1}]/\sum_{j=1}^{n} \epsilon_j .$$

Then it follows from the individual ergodic theorem that W_n converges with probability one to $P\{X_1 = 1, X_2 = 1\}$, provided $P\{Y_1 = 1\} > 0$.

Clearly we may use the same technique to estimate all joint probabilities of the $\{X_n\}$ process. Thus our problem is to give conditions which imply the ergodicity of the $\{Y_n, Z_n\}$ process. In Section 3 we shall assume that the $\{Y_n\}$ process consists of independent random variables. In Section 4 we assume a weaker hypothesis on the $\{Y_n\}$ process but a correspondingly stronger one on the $\{X_n\}$ process.

2. Stationarity of the $\{Y_n, Z_n\}$ process. For $i = 1, 2$ let Ω_i be a set, $\mathfrak{F}_i$ a σ-algebra of subsets of Ω_i, and P_i a probability measure defined on $\mathfrak{F}_i$. Further let $(\Omega, \mathfrak{F}, P)$ be the cross product probability space of $(\Omega_1, \mathfrak{F}_1, P_1)$ and $(\Omega_2, \mathfrak{F}_2, P_2)$, i.e. $\Omega = \Omega_1 \times \Omega_2$, $\mathfrak{F} = \mathfrak{F}_1 \times \mathfrak{F}_2$, and $P = P_1 \times P_2$.

Now suppose there is a point transformation S mapping Ω_2 into itself which is measurable and measure preserving, i.e. if $A_2 \,\varepsilon\, \mathfrak{F}_2$ then $S(A_2) \,\varepsilon\, \mathfrak{F}_2$ and $P_2[S(A_2)] = P_2[A_2]$. Suppose also that for each $\omega_2 \,\varepsilon\, \Omega_2$, there is a measurable and measure preserving point transformation S_{ω_2} mapping Ω_1 into itself. Let T be be the point transformation defined on Ω by $T(\omega_1, \omega_2) = (S_{\omega_2}(\omega_1), S(\omega_2))$, and assume T is measureable, i.e. if $A \,\varepsilon\, \mathfrak{F}$ then $TA \,\varepsilon\, \mathfrak{F}$. Suppose further that

$$[S(\omega_2') = S(\omega_2'')] \Rightarrow S_{\omega_2'} = S_{\omega_2''} \quad \text{a.e. } P_1$$

for almost all $\omega_2'[P_2]$.

LEMMA. *T is measure preserving,* i.e. $P(TA) = P(A)$ *for* $A \,\varepsilon\, \mathfrak{F}$.

PROOF. Since $\mathfrak{F}$ is generated by sets of the form $A = A_1 \times A_2$ with $A_i \,\varepsilon\, \mathfrak{F}_i$ it is sufficient to prove the lemma for such sets. But, using Fubini's Theorem,

$$P[T(A_1 \times A_2)] = \iint_{T(A_1 \times A_2)} dP_1(\omega_1)\, dP_2(\omega_2)$$
$$= \iint_{\{(\omega_1,\omega_2):\, \omega_1 = S_{\omega_2^*}(\omega_1^*),\, \omega_2 = S(\omega_2^*),\, \omega_i^* \varepsilon A_i\}} dP_1(\omega_1)\, dP_2(\omega_2)$$

$$= \int_{S(A_2)} \left[\int_{\{\omega_1 : \omega_1 = S_{\omega_2^*}(\omega_1^*) \text{ for all } \omega_2^* \text{ with } \omega_2 = S(\omega_2^*), \omega_1^* \varepsilon A_1\}} dP_1(\omega_1) \right] dP_2(\omega_2)$$

$$= \int_{S(A_2)} [P_1\{S_{\omega_2^*}(A_1)\}]\, dP_2(\omega_2) \quad \text{where} \quad S(\omega_2^*) = \omega_2$$

$$= \int_{S(A_2)} P_1(A_1)\, dP_2(\omega_2) = P_1(A_1)P_2(SA_2) = P_1(A_1)P_2(A_2). \qquad \blacksquare$$

3. Independent $\{Y_n\}$ process. Let $(\Omega_i, \mathfrak{F}_i, P_i)$, $i = 1, 2$, be as above, but suppose $(\Omega_2, \mathfrak{F}_2, P_2)$ has the following structure. There exists a probability space $(M, \mathfrak{A}, m)$ and $(M_n, \mathfrak{A}_n, m_n)$, $n = 0, \pm 1, \pm 2, \cdots$ are copies of $(M, \mathfrak{A}, m)$. Let

$$\Omega_2 = \bigtimes_{n=-\infty}^{\infty} M_n, \qquad \mathfrak{F}_2 = \bigtimes_{n=-\infty}^{\infty} \mathfrak{A}_n, \quad \text{and} \quad P_2 = \bigtimes_{n=-\infty}^{\infty} m_n.$$

If $\omega_2 \varepsilon \Omega_2$, $\omega_2 = (\cdots, y_{-1}, y_0, y_1, \cdots)$ with $y_i \varepsilon M_i$ for all i, we let $S(\omega_2) = (\cdots, y_0, y_1, y_2, \cdots)$, i.e. S is the shift operator defined on Ω_2. In that case S is one-one onto, bimeasurable, measure preserving, and ergodic. Now suppose for each $\omega_2 \varepsilon \Omega_2$ there exists a point mapping S_{ω_2} of Ω_1 into itself which is bimeasurable and measure preserving. Just as above we define a mapping T of Ω into itself by $T(\omega_1, \omega_2) = S_{\omega_2}(\omega_1), S(\omega_2)$. Assuming that T is measurable it then follows from the lemma that T is measure preserving. Following Kakutani [1], we shall say that a set $A_1 \varepsilon \mathfrak{F}_1$ is invariant with respect to the family $\{S_{\omega_2}, \omega_2 \varepsilon \Omega_2\}$ if $P_1(A_1 \Delta A_{\omega_2} A_1) = 0$ for almost all $\omega_2 \varepsilon \Omega_2$, where $A \Delta B$ is the symmetric difference of A and B, and we shall say that the family $\{S_{\omega_2}, \omega_2 \varepsilon \Omega_2\}$ is ergodic if every $A_1 \varepsilon \mathfrak{F}_1$ which is invariant with respect to the family $\{S_{\omega_2}, \omega_2 \varepsilon \Omega_2\}$ has probability zero or one. The following theorem is proved in [1].

THEOREM 1. (Kakutani). *Suppose for $\omega_2 = (\cdots, y_{-1}, y_0, y_1, \cdots)$, S_{ω_2} depends only on y_0. If the family $\{S_{\omega_2} : \omega_2 \varepsilon \Omega_2\}$ is ergodic then the transformation T is ergodic.*

Now let us return to our $\{X_n\}$ process discussed in the introduction. We may and shall assume that the process is embedded in a process $\{X_n\}$ with n running through all integers. Let $\{Y_n, n = 0, \pm 1, \pm 2, \cdots\}$ be a sequence of positive integer valued random variables such that the $\{Y_n\}$ process is independent of the $\{X_n\}$ process. Let S_1 be the shift operator defined on the $\{X_n\}$ process and let $S_k = S_1^k$ for every integer k. Further let S be the shift operator defined on the $\{Y_n\}$ process. Now assume that the $\{Y_n\}$ process consists of independent, identically distributed random variables, and let $(\Omega_1, \mathfrak{F}_1, P_1)$ be the probability space generated by the $\{X_n\}$ process and $(\Omega_2, \mathfrak{F}_2, P_2)$ the probability space generated by the $\{Y_n\}$ process. Then each $(\omega_1, \omega_2) \varepsilon \Omega$ consists of two infinite sequences, $\omega_1 = (\cdots, x_{-1}, x_0, x_1, \cdots)$, $\omega_2 = (\cdots, y_{-1}, y_0, y_1, \cdots)$. Now define $T(\omega_1, \omega_2) = (S_{y_0}(\omega_1), S(\omega_2))$. Since Y_0 is a discrete random variable it follows that T is measurable and hence we may apply Theorem 1. But note that when we apply the transformation T to the $\{X_n, Y_n\}$ process we get precisely the $\{Y_n, Z_n\}$ process of Section 1. Thus we have

THEOREM 2. *If* (a) *the* $\{X_n\}$ *process is stationary and ergodic, and* (b) *the* $\{Y_n\}$ *process is independent of the* $\{X_n\}$ *process and consists of independent, identically distributed random variables which are positive and integer valued, and* (c) $P\{Y_1 = 1\} > 0$, *then the* $\{Y_n, Z_n\}$ *process is ergodic and we can estimate consistently all joint probabilities of the* $\{X_n\}$ *process by means of observation of the* $\{Y_n, Z_n\}$ *process.*

4. Dependent $\{Y_n\}$ process. In this section we take the probability spaces $(\Omega_i, \mathfrak{F}_i, P_i)$ as in the previous section, except that instead of assuming that P_2 is a product measure and S the shift operator, we merely assume that S is measure preserving and ergodic with respect to P_2. As before suppose for each $\omega_2 \varepsilon \Omega_2$ we have a measure preserving transformation S_{ω_2} of Ω_1 into itself. We assume here also that

$$[S(\omega_2') = S(\omega_2'')] \Rightarrow S_{\omega_2'} = S_{\omega_2''} \quad \text{a.e. } P_1$$

for almost all $\omega_2'[P_2]$. We shall say that the family $\{S_{\omega_2}, \omega_2 \varepsilon \Omega_2\}$ is uniformly strongly mixing if for each $A_1, B_1 \varepsilon \mathfrak{F}_1$ and each $\epsilon > 0$ there exists a positive integer n_0 such that for $n \geqq n_0$ we have

$$|P_1[A_1 \cap S_{S^{n-1}(\omega_2)} S_{S^{n-2}(\omega_2)} \cdots S_{S(\omega_2)} S_{\omega_2} B_1] - P_1(A_1)P_1(B_1)| < \epsilon$$

uniformly in an ω_2-set of probability one.

Defining T as before we have

THEOREM 3. *If the family* $\{S_{\omega_2}, \omega_2 \varepsilon \Omega_2\}$ *is uniformly strongly mixing then* T *is ergodic.*

PROOF. It will be sufficient to show that

$$\lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} P\{A \cap T^k B\} = P(A)P(B)$$

for sets of the form $A = A_1 \times A_2$, $B = B_1 \times B_2$ with $A_i, B_i \varepsilon \mathfrak{F}_i$.

For such sets we find, after some computation, that

$$P\{A_1 \times A_2 \cap T^k(B_1 \times B_2)\}$$
$$= \int_{A_2 \cap S^k B_2} P_1\{A_1 \cap S_{S^{k-1}(\omega_2)} S_{S^{k-2}(\omega_2)} \cdots S_{\omega_2} B_1\} \, dP_2(\omega_2).$$

Hence

$$\lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} P\{A_1 \times A_2 \cap T^k(B_1 \times B_2)\}$$
$$= \lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} P_1(A_1)P_1(B_1)P_2(A_2 \cap S^k B_2) = P_1(A_1)P_1(B_1)P_2(A_2)P_2(B_2)$$

since S is ergodic.

If we apply this result to the problem at hand we obtain

THEOREM 4. *If* (a) *the* $\{X_n\}$ *process is stationary and mixing, and* (b) *the* $\{Y_n\}$ *process is independent of the* $\{X_n\}$ *process and is a stationary, ergodic process with positive integer valued random variables, and* (c) $P\{Y_1 = 1, \cdots, Y_m = 1\} > 0$ *for every positive integer* m, *then the* $\{Y_n, Z_n\}$ *process is ergodic and we may estimate consistently all joint probabilities of the* $\{X_n\}$ *process from observation of the* $\{Y_n, Z_n\}$ *process.*

In this paper we have considered only discrete parameter processes. However the case of a continuous parameter process $X(t)$ is likely to be of considerably greater interest, since in practice such processes are always observed via discrete sampling schemes. We believe that the methods of this paper will apply to such processes and hope to consider them in the near future.

REFERENCE

[1] KAKUTANI, S. (1951). Random ergodic theorems and Markoff processes with a stable distribution. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 247–261. Univ. of California Press.