Z. Wahrscheinlichkeitstheorie verw. Geb. 5, 180—186 (1966)

On Certain Applications of the Spectral Representation of Stationary Processes

R. M. LOYNES

Received December 18, 1964

1. Introduction

It has been known for a long time that a second-order stationary (or covariance-stationary) stochastic process has a spectral representation as a Fourier integral defined in a suitable sense; the covariance function itself has also a Fourier integral representation. (See DOOB [2], Chapters 10 and 11, and the Appendix.) Recently (LOYNES [7]) it has been shown that the notion of second-order stationarity can be usefully generalised, in such a way that a spectral representation again exists.

Yet in spite of this long history, there has rarely seemed to be any connection between the second-order properties and other aspects of stochastic processes, except in the Gaussian case.

In this paper our aim is to show that second-order theory can on occasion be usefully applied to situations to which at first sight it has no relevance. Here two results are obtained: the existence of a Fourier representation for the transition probabilities of certain discrete-time stationary Markov chains, and the existence of a limit connected with the distributions of processes with stationary increments. The first result was originally obtained by KENDALL [4], even with fewer restrictions, but in an entirely different way; the second appears to be new.

2. Fourier Representations for Transition Probabilities

Suppose first that $\{X_n : n \geqq 1\}$ or $\{X_n : -\infty < n < \infty\}$ is a strictly stationary stochastic process, with a denumerable set of possible values which we label $1, 2, 3, \ldots$ for convenience. Let $\mathfrak{B}_n = \mathfrak{B}(X_n, X_{n-1}, \ldots)$ be the σ-field generated by the random variables indicated.

Theorem 1. *We have*

$$P[X_m = j, X_n = k] = \int_{-\pi}^{\pi} e^{i(n-m)\theta}\, d\mu_{jk}(\theta),$$

where $\mu_{jk}(\theta)$ is a function of bounded variation on $[-\pi, \pi]$, and for each j, $\mu_{jj}(\theta)$ is non-decreasing. Consequently provided $p_j \equiv P[X_m = j] \neq 0$,

$$P[X_n = k \mid X_m = j] = p_j^{-1} \int_{-\pi}^{\pi} e^{i(n-m)\theta}\, d\mu_{jk}(\theta).$$

If $\mathfrak{B}_{-\infty} \equiv \bigcap_{n=-\infty}^{\infty} \mathfrak{B}_n$ is a σ-field containing only sets of probability zero and one, then the functions $\mu_{jk}(\theta)$ are absolutely continuous except at $\theta = 0$.

Here and elsewhere there is some freedom in the treatment of the end-points of the range of integration. For definiteness we may suppose the lower end-point included and the upper one excluded.

The proof is very simple. Let Y_n be the random vector with components

$$Y_{n,1} = I(X_n = 1), \ Y_{n,2} = I(X_n = 2), \dots$$

where $I(\cdot)$ is the indicator function of the event indicated. Then $\{Y_n\}$ is a multivariate second-order stationary process, with covariances

$$E(Y_{m,j}\, Y_{n,k}) = E[I(X_m = j)\, I(X_n = k)] = P[X_m = j \cap X_n = k],$$

and now the representation in the theorem is a standard result (Doob [2], p. 596). The simplest way to prove absolute continuity under the stated condition is to use the Wold decomposition (Doob, p. 576). The deterministic part of Y_n is clearly measurable with respect to $\mathfrak{B}_{-\infty}$, and is in consequence constant with probability one. Hence the $\mu_{jk}(\theta)$, which are the spectra and cross-spectra of the Y_n process, can differ from the spectra and cross-spectra of the completely non-deterministic part of Y_n only because of a jump at the origin, and are therefore absolutely continuous except at that point.

Remarks: (i) The condition on $\mathfrak{B}_{-\infty}$ has been used by Rosenblatt [8], who has called such processes completely non-deterministic.

(ii) More complicated probabilities can be expressed in a rather similar way, by using the process $\{Y_n\}$, since for example

$$P[X_l = j, X_m = k, X_n = p] = E[Y_{l,j}\, Y_{m,k}\, Y_{n,p}].$$

(Cf. Blanc-Lapierre and Fortet [1], p. 427, for expectations of such triple products.) In principle all the properties of $\{X_n\}$ can be expressed in terms of the measures which arise thus, although in practise this will of course be difficult.

(iii) It would be interesting to know what functions $\mu_{jk}(\theta)$ are possible; in particular what are the possibilities for $\mu_{jj}(\theta)$, which is the spectral distribution function of a process taking only two values 0 and 1. It would be equally interesting in the case when $\{X_n\}$ is Markov.

(iv) We shall always restrict ourselves to discrete state-spaces, but analogous results hold in other cases. We could for example deal with $P[X_m \in A, X_n \in B]$ for all measurable A and B in the state-space.

(v) Throughout the paper, attention has for simplicity been restricted to processes in discrete time. There is, however, no difficulty in obtaining corresponding results for continuous-time processes.

By far the most interesting special case is that of Markovian $\{X_n\}$, since the transition probabilities are then precisely what are required for the study of the process.

If the chain contains a positive class we can suppose it to be in equilibrium with the associated stationary distribution, and then if j is in this positive class $p_j > 0$, and so the transition probability $p_{jk}^{(n)}$ is given as a Fourier integral. Furthermore it is quite easy to show that if a chain contains only a single positive class, which is aperiodic, then $\mathfrak{B}_{-\infty}$ is trivial: hence the measures μ_{jk} are absolutely continuous except at the origin. Now if the class has instead period $d > 1$, we need only consider the chain $\{X_{nd}\}$ to obtain the final conclusion:

The transition probabilities in a positive-recurrent irreducible Markov chain have Fourier representations

$$p_{jk}^{(n)} = p_j^{-1} \int_{-\pi}^{\pi} e^{in\theta}\, d\mu_{jk}(\theta) \qquad (n \geqq 0),$$

where the functions $\mu_{jk}(\theta)$ are absolutely continuous save for jumps at integral multiples of $\pm 2\pi/d$. (It is clearly easy to identify the jumps of $\mu_{jk}(\theta)$.)

This result is of course merely a restatement of the special case of KENDALL's Theorem III [4] under the hypothesis of positive recurrence. Indeed KENDALL goes slightly further, obtaining the derivative of $\mu_{jj}(\theta)$ explicitly.

In any case however this is only a part of KENDALL's theorem, since he needs no restriction to positive-recurrence, and it would be nice to be able to drop this restriction in the present approach. It has not been found possible to drop all restrictions, but two proofs for the case in which a stationary measure exists will be given; in particular all recurrent chains will then be included.

Suppose then that the chain has an invariant measure m_i, so that $\sum\limits_i m_i p_{ij} = m_j$, with $m_i > 0$ for all i.

One approach is to set up a generalised stationary stochastic process $\{X_n\}$, so that for example

$$m(X_n = i) = m_i,\; m(X_n = i,\; X_q = j) = m_i\, p_{ij}^{(q-n)}$$

if $q \geqq n$. The measure $m(\cdot)$ then takes the place of the probability measure and may be infinite for certain events. However the process $\{Y_n\}$ can be defined in precisely the same way as before, and the proof of the existence of Fourier representations remains valid.

Alternatively there is a slightly different proof which avoids the use of these generalised stochastic processes. Let $\{X_n\}$ be the Markov chain starting from an arbitrary initial state i, and consider a particular component of Y_n, say $Y_{n,j}$. Neither of these processes are stationary, but nevertheless the covariance function of $Y_{n,j}$ is positive-definite. (A function $f(m, n)$ of two variables is positive-definite if $\sum f(m, n)\, z_m \bar{z}_n \geqq 0$ for all finite sets of complex numbers z_n.) The proof consists in observing that $E\,|\sum z_n Y_{n,j}|^2 \geqq 0$; furthermore except in the trivial case when $Y_{n,j}$ is uniquely determined by $Y_{n-1,j}$ there is strict positive-definiteness.

We therefore know that

$$E[Y_{m,j}\, Y_{n,j}] = p_{ij}^{(m\wedge n)}\, p_{jj}^{(|m-n|)}$$

is positive-definite (as a function of m and n), where $m \wedge n = \min(m, n)$. Multiplying by m_i and summing over i, it follows that $p_{jj}^{(|m-n|)}$ is (strictly) positive-definite, and consequently by the theorem of Herglotz it has a Fourier representation. With a little complication we can also deal with $p_{jk}^{(n)}$ for $j \neq k$.

There is also another method (due to J. F. C. KINGMAN) which is only valid for recurrent chains, but which has the advantage that it could be applied to certain non-Markovian (non-stationary) processes.

We have just seen that

$$p_{jj}^{(m\wedge n)}\, p_{jj}^{(|m-n|)}$$

is positive-definite. Similarly so is

$$p_{jj}^{(m \wedge n+p)} \; p_{jj}^{(|m-n|)}$$

for any $p \geqq 0$. On summing these from $p = 1$ to $p = t$, dividing by $\sum_{1}^{t} p_{jj}^{(p)}$, and then allowing t to tend to ∞, it follows since the process is recurrent that

$$p_{jj}^{(|m-n|)}$$

is positive-definite.

Returning now to the general problem of obtaining Fourier representations for various quantities, we conclude by showing that certain non-stationary processes satisfying different hypotheses give rise to results of this type.

Suppose that $P[X_m = X_n]$ is a function of $n - m$ only; this would be the case for example if $\{X_n\}$ has stationary increments. Then it is possible to show that

$$P[X_m = X_n] = \int_{-\pi}^{\pi} e^{i(n-m)\theta} \, d\mu(\theta),$$

where $\mu(\theta)$ is a bounded monotone function. This again can be proved by using the process $\{Y_n\}$ defined earlier, since

$$P[X_m = X_n] = \sum_j E(Y_{m,j} \, Y_{n,j})$$

is clearly positive-definite (strictly positive-definite unless X_n is completely determined by $X_{n-1}, X_{n-2}, \ldots$). For a slightly different proof we define the "generalised covariance" of Y_n as

$$\sum_j E(Y_{m,j} \, Y_{n,j}),$$

so that $\{Y_n\}$ is a "generalised" second-order stationary stochastic process, and the result follows by Theorem 1 of [7].

3. Stationary Sequences of Random Operators between Hilbert Spaces

The next section will be concerned with processes with stationary increments taking values in a locally compact topological group (including of course as a special case real-valued processes). The method we use to treat such processes is to apply the theory of generalised second-order stationary stochastic processes developed in [7] to the group representations, and we have for clarity and possible future applications separated that part of the theory which applies to more general operators in Hilbert space into this section. The ideas and terminology of [7] will be used freely.

Let H_1 and H_2 be two separable Hilbert spaces; the separability is not altogether necessary, but allows us to avoid certain measurability difficulties.

Let $\mathfrak{B}(H_1, H_2)$ be the (Banach) space of bounded linear operators from H_1 to H_2. If $A \in \mathfrak{B}(H_1, H_2)$, define the *adjoint* of A, A^*, by

$$(Ax, y) = (x, A^* y) \qquad x \in H_1, \; y \in H_2.$$

Then $A^* \in \mathfrak{B}(H_2, H_1)$. (Note that the two inner products in the preceding equation are defined on different spaces.)

We shall show that $\mathfrak{B}(H_1, H_2)$ can be made an LVH-space (see [6], and § 2 of [7]). To do this we must define a strongly admissible space Z, and a suitable map from the Cartesian product of $\mathfrak{B}(H_1, H_2)$ with itself to Z. We therefore define Z to be $\mathfrak{B}(H_1) \equiv \mathfrak{B}(H_1, H_1)$ with the weak topology and write for the vector inner-product of $A, B \in \mathfrak{B}(H_1, H_2)$

$$[A, B] = B^* A.$$

The fact that this space Z is strongly admissible was observed at the end of [6].

The generalised random variables which will form the subject of the following discussion take their values in $\mathfrak{B}(H_1, H_2)$. We therefore define a (generalised) random variable to be a function X from some fixed probability space to $\mathfrak{B}(H_1, H_2)$ which is weakly measurable: i.e. for each $x \in H_1, y \in H_2$, $(X x, y)$ is a scalar measurable function (an ordinary random variable).

We shall show next that Assumptions 1 to 4 of § 3 of [7] hold in the present context, and it will then be possible to apply the theory described there. It will be observed that the situation treated here is merely a natural generalisation of that treated in § 4 of [7], where the random variables were $p \times q$ matrices.

Assumption 1 is trivially true. The induced random variables appearing in Assumption 2 are easily shown to be weakly measurable, and it is therefore sufficient to define the expectation of (a subclass of the) weakly measurable random variables. Let W be one: then if for each $x, y \in H_1$ $E(Wx, y)$ is well-defined and its modulus is bounded above by $K|x|\,|y|$ for some finite positive constant K we define EW by

$$(EWx, y) = E(Wx, y)$$

It follows at once that Assumptions 2 and 3 are satisfied. The proof that Assumption 4 is satisfied here can be carried out, using the Riesz-Fischer theorem, in a way similar to the proof of the completeness of the sequence space l_2.

It is also possible to describe the null random variables. By virtue of the separability, it is easy to see that $E[X, X] = E[X^* X] = 0$ if and only if $X = 0$ with probability one.

Now that we have verified that the assumptions are valid, the results may be applied. If for example $\{X_n : -\infty < n < \infty\}$ is a generalised second-order stationary process, i.e. if $E[X_n, X_m] \equiv E X_m^* X_n$ is a function of $n - m$ only, X_n and the covariance-function both have Fourier representations, and ergodic theorems are valid.

4. A Property of Processes with Stationary Increments

We shall prove the following theorem, concerning random variables which take their values in a separable locally compact topological group G.

By a random variable we mean, of course, a function from some probability space to G, which is measurable with respect to the Borel sets of G, and the probability distribution of a random variable X is the measure μ_X defined on the Borel sets of G by

$$\mu_X(A) = P[X \in A].$$

A sequence of distributions μ_n converges vaguely to a measure ν on G if

$$\int f(g)\,\mu_n(dg) \to \int f(g)\,\nu(dg)$$

for all continuous scalar-valued functions f, which vanish outside a compact set, and the sequence converges weakly to ν if this relationship is valid for all bounded continuous functions f. In the latter case ν also is a probability distribution. Vague and weak convergence coincide if G is compact. (See GRENANDER [3], especially § 2.1, for these concepts.)

Theorem 2. *Let $\{X_n : n \geqq 1\}$ be a stochastic process taking values in G with the property that the distribution of $X_m^{-1} X_n$ depends only on $n - m$. Then*

$$\frac{1}{n}(\mu_{X_1} + \mu_{X_2} + \cdots + \mu_{X_n})$$

converges vaguely to some measure ν as $n \to \infty$; there is weak convergence to a probability distribution if G is compact.

Remarks: (i) The hypothesis of the theorem is actually weaker than that of stationary increments.

(ii) Because of the symmetry between left and right products in the group, the result would also be true if the distribution of $X_n X_m^{-1}$ depends only on $n - m$.

(iii) The theorem contains an earlier result due to GRENANDER ([3], Theorem 3.0), except of course for the evaluation of the limit.

(iv) There are similarities between the present theorem and Theorem 6 of [7].

(v) Both statement and proof of the theorem are simpler in the case when G is the real line. Then vague convergence of distributions is equivalent to convergence of distribution functions at continuity points of the limit, and the convergence is weak if the limit is an honest distribution function. The proof merely involves consideration of the random variables $\exp\{i\,\theta\,X_n\}$ for fixed but arbitrary θ, which under the stated hypothesis obviously forms an ordinary second-order stationary process: the result then follows easily. Consideration of this example also suggests that in stating the result of the theorem we have not exhausted all the information that this approach can give.

Proof. There exists a complete set of irreducible representations of G as unitary operators in Hilbert space. Let $U(g)$ $(g \in G)$ be a particular representation, of this complete set, acting in the Hilbert space H: because of the separability of G, H is also separable. Now as a function of g, $U(g)$ is weakly continuous, and it follows that $U(X_n)$ is a weakly measurable function on the probability space for each n, so that it is a generalised random variable of the type considered in the last section. Furthermore

$$E[U(X_n), U(X_m)] = E\,U(X_m)^*\,U(X_n) = E\,U(X_m^{-1})\,U(X_n) = E\,U(X_m^{-1}X_n)$$

using the properties of the representation, and by hypothesis this is a function of $n - m$ only, so that the sequence $\{U(X_n)\}$ is (generalised) second-order stationary.

It follows from the one-sided parameter set version of Theorem 3 of [7] that

$$\frac{1}{n}\{U(X_1) + U(X_2) + \cdots + U(X_n)\}$$

second-order converges to some limit as $n \to \infty$, and on taking expectations it follows that

$$\frac{1}{n}\{E\,U(X_1) + E\,U(X_2) + \cdots + E\,U(X_n)\}$$

converges weakly.

Now $E\,U(X_r)$ as a function of the representation U is just what has been called the Fourier transform of the distribution of X_r, in [3] § 5.2 and [5], and we have therefore just proved that the Fourier transforms of the measures

$$\frac{1}{n}\{\mu_{X_1} + \mu_{X_2} + \cdots + \mu_{X_n}\}$$

converge weakly as $n \to \infty$. From Theorem 1 of [5] the final result follows.

References

[1] BLANC-LAPIERRE, A., and R. FORTET: Théorie des fonctions aléatoires. Paris: Masson 1953.
[2] DOOB, J. L.: Stochastic processes. New York: Wiley 1954.
[3] GRENANDER, U.: Probabilities on algebraic structures. New York: Wiley and Sons 1963.
[4] KENDALL, D. G.: Unitary dilations of Markov transition operators. In Surveys in probability and statistics. Stockholm: Almquist and Wiksell 1959.
[5] LOYNES, R. M.: Fourier transforms and probability theory. Ark. Mat. **5**, 37—42 (1963).
[6] — Linear operators in VH-spaces. Trans. Amer. math. Soc. **116**, 167—180 (1965).
[7] — On a generalization of second-order stationarity. Proc. London math. Soc., III. Sev. **15**, 385—398 (1965).
[8] ROSENBLATT, M.: Stationary processes as shifts of functions of independent random variables. J. Math. Mech. 8, 665—682 (1959).

Statistical Laboratory,
University of Manchester