NOTES ON CONTINUOUS STOCHASTIC PHENOMENA

By P. A. P. MORAN, *Institute of Statistics, Oxford University*

The study of stochastic processes has naturally led to the consideration of stochastic phenomena which are distributed in space of two or more dimensions. Such investigations are, for instance, of practical interest in connexion with problems concerning the distribution of soil fertility over a field or the relations between the velocities at different points in a turbulent fluid. A review of such work with many references has recently been given by Ghosh (1949) (see also Matérn, 1947). In the present note I consider two problems arising in the two- and three-dimensional cases.

Relations between continuous and discontinuous processes

Stochastic variables defined for points on a plane may be considered as defined at a discrete set of points (for example, at all points with integral co-ordinates) or for a continuous domain of points. The latter is the natural approach when considering soil fertility, but in the study of the efficiency of experimental designs it is more natural to consider the fertility as varying discontinuously from plot to plot rather than within each plot. For this reason I begin by considering the relationship between continuous and discrete models of such phenomena.

First consider stationary stochastic processes in one dimension defined by variates $x(t)$, where t is 'time' and takes either integral or a continuous range of values. Continuous processes whose variate $x(t)$ has a correlation function

$$\rho(t) = \exp[-\lambda\,|\,t\,|] \tag{1}$$

are known (Bartlett, 1947, p. 79) to exist and to have a spectral density given by

$$W'(\theta) = \frac{2\lambda}{\pi(\lambda^2+\theta^2)}, \tag{2}$$

so that
$$\rho(t) = \int_0^\infty \cos t\theta\, dW(\theta) = 2\lambda \int_0^\infty \frac{\cos t\theta\, d\theta}{\pi(\lambda^2+\theta^2)}.$$

From such a continuous process, a discrete process can be derived in two ways. First we might consider the values of $x(t)$ only at discrete values of t ($= 0, \pm 1, \ldots$ say). Such a process would have the serial correlation

$$\rho_s = \exp[-\lambda\,|\,s\,|] \quad (s = 0, \pm 1, \ldots),$$

and could be regarded as being generated by a simple Markoff relation of the form

$$x_s = e^{-\lambda} x_{s-1} + \eta_s,$$

where $\{\eta_s\}$ is a stationary process which is not necessarily completely random but nevertheless has all its serial correlations zero.

In practice it is perhaps more realistic to consider discrete processes derived from continuous ones in another way. Suppose we write

$$X(s) = \int_s^{s+1} x(t)\, dt, \tag{3}$$

where s takes values 0, ± 1, ... and the integral is a 'stochastic integral'. A completely rigorous discussion of this definition would be neither short nor (in the present connexion) very illuminating, and we do not enter into such considerations here. (For a general discussion of such questions see Lévy, 1948.) It is clear that $\rho(t)$ is continuous at $t = 0$ and consequently $x(t)$ is 'stochastically continuous' (Bartlett, 1947, p. 77). To discuss the correlational and spectral properties of $X(s)$ we may therefore argue as follows. We approximate to

$$X(s) = \int_s^{s+1} x(t)\,dt$$

by a sum, and to find $\operatorname{var} X(s)$ and $\operatorname{cov}\{X(s), X(s+k)\}$, we take the expectations of the sums and proceed to the limit. In this way we find

$$\begin{aligned}\operatorname{var} X(s) = E\{X(s)\}^2 &= \int_0^1 da \int_0^1 db \exp[-\lambda\,|\,a-b\,|] \\ &= \frac{2}{\lambda^2}\{\lambda - 1 + e^{-\lambda}\}.\end{aligned}$$

Similarly, for $k \geqslant 1$

$$\begin{aligned}E\{X(s)\,X(s-k)\} = E\{X(s)\,X(s+k)\} &= \int_0^1 da \int_0^1 db \exp[-\lambda\,|\,k+b-a\,|] \\ &= \frac{1}{\lambda^2} e^{-\lambda k}(e^{\lambda}-1)(1-e^{-\lambda}).\end{aligned}$$

Thus ρ_k ($k \geqslant 1$), the serial correlation of $X(s)$ and $X(s+k)$, is given by

$$\rho_k = \rho_{-k} = \frac{e^{-\lambda k}(e^{\lambda}-1)(1-e^{-\lambda})}{2(\lambda - 1 + e^{-\lambda})} = A\,e^{-\lambda k}, \quad \text{say.} \tag{4}$$

Then as $\lambda \to 0$, $\rho_k \to 1$ and as $\lambda \to \infty$, $\rho_k \to 0$ as we expect. Now, using the fact that

$$\sinh\lambda = \tfrac{1}{2}(e^{\lambda} - e^{-\lambda}) > \lambda \quad \text{for} \quad \lambda > 0,$$

it can be easily shown that

$$(e^{-\lambda}-1)(1-e^{-\lambda}) > 2(\lambda - 1 + e^{-\lambda}),$$

and therefore

$$A > 1.$$

On the other hand, it is easily verified algebraically that $\rho_1 = A\,e^{-\lambda} < 1$. Formula (4) is not the sort of correlation function which would be obtained for a process which is the solution of a simple stochastic difference equation of Markoff type because $A > 1$, but we may construct a simple mechanism which would generate a process having the above correlational properties. Write $r = e^{-\lambda}$. Then the serial correlation generating function of the process $X(s)$ is

$$\begin{aligned}\sum_{-\infty}^{\infty} \rho_k z^k &= 1 + A\sum_1^{\infty} r^k z^k + A\sum_1^{\infty} r^k z^{-k} \\ &= 1 + \frac{Arz}{1-rz} + \frac{Arz^{-1}}{1-rz^{-1}} \\ &= \frac{1 + r^2 - 2Ar^2 + (A-1)\,rz + (A-1)\,rz^{-1}}{(1-rz)(1-rz^{-1})} \\ &= \frac{(\alpha+\beta z)(\alpha+\beta z^{-1})}{(1-rz)(1-rz^{-1})},\end{aligned}$$

where

$$\alpha, \beta = \tfrac{1}{2}\{(1-r)^2 + 2Ar(1-r)\}^{\frac{1}{2}} \pm \tfrac{1}{2}\{(1+r)^2 - 2Ar(1+r)\}^{\frac{1}{2}}.$$

α and β can be easily verified algebraically to be real. Then if $\{\xi_n\}$ is a sequence of independent random variables with zero means and the same standard deviation, the process $\{Y_n\}$ generated by the relation

$$Y_{n+1} = rY_n + \alpha\xi_n + \beta\xi_{n-1},$$

will have the same correlational properties as $\{X_n\}$.

We now consider what happens when we generalize the above to 'processes' with more than one 'time'. As the idea of a continuing 'process' is now less applicable we shall call such a model a 'spatial stochastic system'. We consider such a system to be defined if for any set of points $P_1, \ldots, P_k$ in such a plane (or higher dimensional space) there is given a set of random variables $x_1, \ldots, x_k$ and a corresponding joint distribution function $F(x_1, \ldots, x_k)$ which satisfies the customary consistency conditions that the joint distribution of any set of x's, $x_1, \ldots, x_p$ $(p<k)$ is obtained from the joint distribution of $x_1, \ldots, x_k$ by integrating out $x_{p+1}, \ldots, x_k$. This condition corresponds to the Chapman-Kolmogoroff equation in the theory of processes with a single 'time'. If the distribution function $F(x_1, \ldots, x_k)$ is invariant under any translation of the set of points $P_1, \ldots, P_k$ we call the system 'stationary', and if, in addition, it is invariant under any rotation we call it 'isotropic'. If we only know that the first- and second-order moments of $x_1, \ldots, x_k$ are invariant under such a translation (or rotation), we say the system is stationary (or isotropic) to the second order. In what follows we consider stationary processes, but the results obviously hold under the weaker condition also.

Suppose the variates defined so that each has zero expectation and variance σ^2. Considering first systems in two dimensions we take two parameters t and u to correspond to the 'times', and we then have a correlation function

$$\rho(p,q) = \text{correlation}\,\{x(t,u),\, x(t+p, u+q)\}.$$

It follows that

$$\rho(p,q) = \rho(-p, -q),$$

but it is not necessarily true that

$$\rho(p,q) = \rho(p, -q) \quad \text{or} \quad \rho(-p,q).$$

The latter relations would be true if the system were isotropic, but they are not a sufficient condition for isotropy.

The natural generalization of processes with a correlational function (1) would be a system whose correlational function is

$$\rho(p,q) = \exp\left[-\lambda(p^2+q^2)^{\frac{1}{2}}\right]. \tag{5}$$

As will be seen later it is easy to show that such systems exist. Now suppose we derive from a continuous system having (5) as correlation function, a two-dimensional discrete system defined by a system of variables X_{lm}, where l, m take integral values, and X_{lm} is defined by

$$X_{l,m} = \int_l^{l+1}\int_m^{m+1} x(t,u)\,dt\,du.$$

Following our previous argument we might expect that we should have a correlation function which generalizes (5), i.e. of the form (5) multiplied by a constant. That this is not true can be seen as follows. If the result were true, the covariance of X_{lm} and $X_{l+p,m+q}$ would be of the form

$$K\exp\left[-\mu(p^2+q^2)^{\frac{1}{2}}\right],$$

where K is some constant and μ is not necessarily equal to λ. It would then follow that

$$T = \exp\left[\mu(p^2+q^2)^{\frac{1}{2}}\right]\operatorname{cov}(X_{l,m}, X_{l+p,m+q})$$

2-2

would be independent of p and q. Taking in particular the case $q = 0$, this is equal to

$$T = \int_0^1 dt_1 \int_0^1 du_1 \int_0^1 dt_2 \int_0^1 du_2 \exp\{\mu p - \lambda[(p+t_2-t_1)^2 + (u_2-u_1)^2]^{\frac{1}{2}}\}.$$

Then
$$\frac{\partial T}{\partial p} = \int_0^1 dt_1 \int_0^1 du_1 \int_0^1 dt_2 \int_0^1 du_2 \exp\{\mu p - \lambda[(p+t_2-t_1)^2 + (u_2-u_1)^2]^{\frac{1}{2}}\}$$
$$\times \left\{\mu - \frac{\lambda(p+t_2-t_1)}{\{(p+t_2-t_1)^2 + (u_2-u_1)^2\}^{\frac{1}{2}}}\right\}.$$

If $\mu \geqslant \lambda$ the second factor is greater than zero over the whole range of integration with the possible exception of the points where $u_2 = u_1$, and so the whole integral is non-zero, which contradicts the hypothesis. On the other hand, if $\mu < \lambda$ we can choose p so large that over the whole range of integration the second factor is negative, again contradicting the hypothesis. Thus this generalization of the previous argument breaks down.

It follows that in cases where it is necessary to derive a discrete system from a continuous one, in the above manner, it would be more convenient, if perhaps occasionally less realistic, to consider systems with a correlation function

$$\rho(p,q) = \exp[-\lambda\,|\,p\,| - \mu\,|\,q\,|]. \tag{6}$$

It is easy to see that in this case the discrete process $X_{l,m}$ has a correlation function

$$\rho_{kl} = A_1 A_2 \exp[-\lambda\,|\,k\,| - \mu\,|\,l\,|],$$

where
$$A_1 = \frac{(e^{\lambda}-1)(1-e^{-\lambda})}{2\{\lambda - 1 + e^{-\lambda}\}},$$

and
$$A_2 = \frac{(e^{\mu}-1)(1-e^{-\mu})}{2\{\mu - 1 + e^{-\mu}\}},$$

when neither k nor l is equal to zero. When one is equal to zero the formula is modified.

Cases where correlation functions of this kind can plausibly arise are those in which we might attempt to study the relative efficiencies of various experimental designs assuming that soil fertility is a random variable with a spatial correlation of the form (6).

The existence of spatial stochastic systems with prescribed correlation functions follows from the two- (or more) dimensional analogue of Khintchine's theorem. This asserts that a necessary and sufficient condition for the existence of a spatial stochastic system with a correlation function $\rho(p,q)$ is that there exist a function $W(x,y)$, non-decreasing in x and y, such that

$$\rho(p,q) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(px+qy)}\,dW(x,y)$$

and
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dW(x,y) = 1.$$

The sufficiency of this condition is easily proved by a simple generalization of Khintchine's method (1934) of constructing a stochastic process whose correlation function is the desired one, whilst the necessity follows from a theorem of H. Cramér (1939) (see also Lévy, 1948). $W(x,y)$ can then be found in terms of $\rho(p,q)$ by a Fourier inversion formula and in particular if the system is isotropic, $W_{xy}(x,y)$, if it exists, will be a function of (x^2+y^2) only.

We can thus prove that $\exp[-\lambda(p^2+q^2)^{\frac{1}{2}}]$ is a possible correlation function (as proved in Matérn, 1947, p. 27). It is clearly sufficient to take $\lambda = 1$. We then verify that $W(x,y)$ can be taken as

$$\frac{1}{2\pi}\int_{-\infty}^{x}\int_{-\infty}^{y} \frac{du\,dr}{(1+u^2+r^2)^{\frac{3}{2}}}.$$

Consider
$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{e^{i(px+qy)}\,dx\,dy}{(1+x^2+y^2)^{\frac{3}{2}}}.$$

Write
$$x = t\cos\alpha,\quad y = t\sin\alpha,\quad p = a\cos\beta,\quad q = a\sin\beta.$$

The above is then equal to
$$\frac{1}{2\pi}\int_0^\infty dt\int_0^{2\pi}d\alpha\,\frac{t\,e^{ita\cos(\alpha-\beta)}}{(1+t^2)^{\frac{3}{2}}}$$
$$= \frac{1}{2\pi}\int_0^\infty dt\int_0^{2\pi}d\alpha\,\frac{t\,e^{ita\cos\alpha}}{(1+t^2)^{\frac{3}{2}}}.$$

Now
$$J_0(ta) = \frac{1}{2\pi}\int_0^{2\pi}e^{ita\cos\alpha}\,d\alpha$$

(Whittaker and Watson, 1935, p. 364), and so the above integral is equal to
$$\int_0^\infty \frac{tJ_0(ta)}{(1+t^2)^{\frac{3}{2}}}\,dt.$$

Now it is known that
$$\int_0^\infty \frac{tJ_0(ta)\,dt}{(1+t^2)^{\frac{3}{2}}} = \frac{a^{\frac{1}{2}}K(a)}{2^{\frac{1}{2}}\Gamma(\frac{3}{2})}.$$

This formula can be obtained by putting $\nu = 0$, $\mu = \frac{1}{2}$ in formula (7·11·6) of Titchmarsh (1937, p. 201). But
$$K_{\frac{1}{2}}(z) = \left(\frac{\pi}{2z}\right)^{\frac{1}{2}}e^{-z},$$

and so the above integral equals $e^{-a} = \exp[-(p^2+q^2)^{\frac{1}{2}}]$. The above method of argument can be easily generalized to three dimensions and $\exp[-\lambda(p^2+q^2+r^2)^{\frac{1}{2}}]$ can be shown to be a possible correlation function. This type of function gives a satisfactory representation of the correlation between velocities in certain cases of turbulence, and it may here be pointed out that the spectral theory of turbulence has recently been developed in a very elegant form by Batchelor (1949).

By using an inversion formula we also see that in two or more dimensions, if the system is isotropic, $W(x, y, \ldots)$ will be a function of $x^2+y^2+\ldots$ only, and by integrating first over all directions, we see that for isotropic processes the correlation function will always be representable as a Fourier-Bessel-Stieltjes integral of the type $\int_0^\infty J_n(ut)\,t\,dF(t)$.

Test of the existence of two dimensional stochastic schemes

Another problem arises in practice when we are given a set of variates $X_{i,j}$ (i, j taking integral values), and we wish to decide whether there is any evidence that these variates are spatially correlated. Such a case can arise, for example, in uniformity trials in agricultural research. This is the two-dimensional analogue of the problem of testing the significance of serial correlation coefficients on which a great deal has been written (for references see Moran, 1948*a*). We give here a simple test for correlation between nearest neighbours which generalizes a method of a previous paper (Moran, 1948*a*).

We suppose that we have mn independent variates x_{ij} ($i = 1, \ldots, m$; $j = 1, \ldots, n$), and we define what seems to be a natural definition of a correlation coefficient between x's which are nearest neighbours. Write
$$mn\bar{x} = \sum_{i,j}x_{ij}\quad\text{and}\quad z_{ij} = x_{ij}-\bar{x},$$

and
$$r_{11} = \left(\frac{mn}{2mn-m-n}\right)\frac{\sum_{i=1}^{m}\sum_{j=1}^{n-1} z_{ij}z_{i,j+1} + \sum_{i=1}^{m-1}\sum_{j=1}^{n} z_{ij}z_{i+1,j}}{\sum_{i,j} z_{ij}^2}$$
$$= \left(\frac{mn}{2mn-m-n}\right)I.$$

The initial factor is conventionally introduced because there are mn terms in the denominator and $2mn-m-n$ in the numerator. In large samples r_{11} could therefore be considered an appropriate estimator of a presumed correlation coefficient between nearest neighbours. As we are here only concerned with a test for randomness it is sufficient to consider a test using I alone. If the x_{ij} are independently distributed in the same distribution, the z_{ij} are all on an equal footing, and if their distribution had a finite second moment, the correlation between any two z's would be $(mn-1)^{-1}$. Without assuming, however, that any moments exist, we have

$$E(I) = E\left\{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n-1} z_{ij}z_{i,j+1} + \sum_{i=1}^{m-1}\sum_{j=1}^{n} z_{ij}z_{i+1,j}}{\sum z_{ij}^2}\right\}$$
$$= (2mn-m-n)\,E\left\{\frac{z_{11}z_{12}}{\sum z_{ij}^2}\right\}$$
$$= \left(\frac{2mn-m-n}{mn(mn-1)}\right)E\left\{\frac{(\sum z_{ij})^2 - \sum z_{ij}^2}{\sum z_{ij}^2}\right\}$$
$$= -\frac{2mn-m-n}{mn(mn-1)},$$

because $\sum z_{ij} = 0$. (The $\sum$ symbol without suffixes is used above to indicate summation over all values of i and j.)

To evaluate the second moment of I we have to assume further that all the x_{ij} are distributed normally. There is then no real restriction in taking the variance of x_{ij} to be unity. The denominator of I is then distributed as χ^2 with $mn-1$ degrees of freedom, and since I is the ratio of a quadratic form in the z_{ij} to a quantity $\sum z_{ij}^2$ distributed as χ^2 it follows that I is itself distributed independently of $\sum z_{ij}^2$. We then have for any positive integer p

$$E(I^p) = \frac{E(\text{numerator of } I)^p}{E(\sum z_{ij}^2)^p}.$$

But from the properties of the χ^2 distribution we have

$$E(\sum z_{ij}^2)^2 = (mn-1)(mn+1).$$

We now have to find the expectation of

$$\left(\sum_{i=1}^{m}\sum_{j=1}^{n-1} z_{ij}z_{i,j+1} + \sum_{i=1}^{m-1}\sum_{j=1}^{n} z_{ij}z_{i+1,j}\right)^2.$$

On multiplying this out we find terms of three different kinds.

(1) Terms of the form $z_{11}^2 z_{12}^2$. The number of such terms is clearly equal to the number of joins of nearest neighbours on the lattice, i.e. to $(2mn-m-n)$. Moreover, by considering the characteristic function of the joint distribution of the z's we see that

$$E(z_{11}^2 z_{12}^2) = (1+2\rho^2)\,\sigma^4,$$

where ρ is the correlation coefficient $(mn-1)^{-1}$ between any two z's and

$$\sigma^2 = \operatorname{var}(z) = (mn-1)/mn.$$

The contribution of terms of this kind to the expectation of the numerator is therefore

$$(2mn-m-n)(1+2(mn-1)^{-2})\sigma^4 = \frac{(2mn-m-n)(m^2n^2-2mn+3)}{(mn-1)^2}\sigma^4.$$

(2) The second type of term is of the form $z_{11}z_{12}^2z_{13}$, and it is easy to see (Moran, 1947, p. 323) that the number of such terms occurring in the square of the numerator is

$$2\{m(n-2)+n(m-2)+4(m-1)(n-1)\} = 4\{3mn-3m-3n+2\}.$$

Moreover, the expectation of such a term is $(\rho+2\rho^2)\sigma^4$, and so the total contribution is

$$-\frac{4(3mn-3m-3n+2)(mn-3)}{(mn-1)^2}\sigma^4.$$

(3) The third type of term is of the form typified by $z_{11}z_{12}z_{33}z_{34}$ and corresponds to two joins on the lattice without common points. The number of such terms in the expansion of the numerator is (Moran, 1947, p. 323)

$$4m^2n^2-4m^2n-4mn^2+m^2+n^2-12mn+13m+13n-8,$$

and the expectation is $3\rho^2\sigma^4$, so that the total contribution is

$$\frac{3}{(mn-1)^2}\{4m^2n^2-4m^2n-4mn^2+m^2+n^2-12mn+13m+13n-8\}\sigma^4.$$

Adding the above contributions and dividing by $(mn-1)(mn+1)$, we find

$$E(I^2) = \frac{2m^3n^3-m^3n^2-m^2n^3-4m^2n^2+2m^2n+2mn^2-2mn+3m^2+3n^2}{m^2n^2(mn-1)(mn+1)},$$

and $\operatorname{var} I$ is best found from the formula

$$\operatorname{var}(I) = E(I^2)-[E(I)]^2.$$

Higher moments of the distribution of I could be calculated in the same way using the frequencies of various combinations of joins given in Moran (1948*b*), but this would be very arduous. If, however, the mean of the distribution of the x's is known exactly, and r_{11} and I defined using deviations from this mean, the formulae are considerably simpler. In both cases it is easy to show that the distribution of I tends to normality as m and n increase. It should also be noticed that a test, based on rearrangements, for randomness in an array of this kind has been given by M. N. Ghosh (1948).

REFERENCES

BARTLETT, M. S. (1947). *North Carolina Lectures on Stochastic Processes.*
BATCHELOR, G. K. (1949). *Proc. Roy. Soc.* A, **195**, 513–32.
CRAMÉR, H. (1939). *Trans. Amer. Math. Soc.* **46**, 191–201.
GHOSH, B. (1949). *Bull. Calcutta Stat. Ass.* **2**, 11–28.
GHOSH, M. N. (1948). *Sankhyā*, **9**, 39–46.
KHINTCHINE, A. (1934). *Math. Ann.* **109**, 604–15.
LÉVY, P. (1948). *Processus Stochastiques.* Paris.
MATÉRN, B. (1947). *Metoder att uppskatta noggrannheten vid linje-och provytetaxering.* Stockholm.
MORAN, P. A. P. (1947). *Proc. Camb. Phil. Soc.* **43**, 321–8.
MORAN, P. A. P. (1948*a*). *Biometrika*, **35**, 255–60.
MORAN, P. A. P. (1948*b*). *J.R. Statist. Soc.* B, **10**, 243–51.
TITCHMARSH, E. C. (1937). *Theory of Fourier Integrals.* Oxford.
WHITTAKER, E. T. & WATSON, G. N. (1935). *Modern Analysis.* Cambridge.