Fontainebleau/CGMM N-858

CHANGE OF SUPPORT IN THE CASE OF DIFFUSION TYPE RANDOM FUNCTIONS

G. MATHERON Octobre 1983

To be published in the "Journal of the International Association for Mathematical Geology

CHANGE OF SUPPORT

IN THE CASE OF DIFFUSION TYPE RANDOM FUNCTIONS

Contents

0. INTRODUCTION 1

1. AN INTRODUCTORY EXAMPLE 3

1.1 The infinitesimal conditional drift 4
1.2 The infinitesimal conditional variogram 4
1.3 A differential equation for the p.d.f. 5
1.4 Moments of order > 2 6
1.5 The law for a small change of support 6

2. DIFFUSION TYPE RANDOM FUNCTIONS 9

2.1 The Ito conditions 9
2.2 A differential equation for the p.d.f. 10
2.3 Structure of the conditional covariance 12
2.4 Variation of the p.d.f. under a change of support 13
2.5 The case of a proportional effect 15

3. EXAMPLES 16

3.1 Diffusion Markov processes 17
3.2 The multigaussian case 19
3.3 The regular case 20

4. APPROXIMATE MODELS OF CHANGE OF SUPPORT 21

4.1 The affine correction 22
4.2 The isofactorial model 24
4.3 The second order approximation 25
4.4 The lognormal case 26
4.5 The gamma diffusion process 29
4.6 The square of the gamma process 32

5. A MORE GENERAL MODEL 32

FIGURES 36

BIBLIOGRAPHY 43

CHANGE OF SUPPORT

IN THE CASE OF DIFFUSION TYPE RANDOM FUNCTIONS

by

G. MATHERON

ABSTRACT

Diffusion type random functions are a first attempt at a non-Markovian and multidimensional generalization of the Ito stochastic integral theory. Within this framework, the variation δf of the p.d.f. can be evaluated for a small change of support. Comparison with the usual approximate models suggests the following conclusions : the affine correction is false for the first order approximation unless z-m is a factor. The isofactorial model is exact for the first order and, in the multigaussian case, is almost correct for the second order approximation. Counter-examples are given in the discontinuous case, and a more general model is suggested.

KEY WORDS : Diffusion processes ; Ito conditions ; conditional drift ; conditional variogram ; change of support.

0. INTRODUCTION

Change of support is probably the most important problem of Geostatistics today. In some parts of the mining industry, the support v of the selection is chosen by taking into account the information available at the time of selection (e.g. following the ore/waste contours). In this case, geostatisticians say that the geometry is adaptative, and only empirical approaches are possible. In what follows, we shall consider only the case of a fixed geometry, where the choice of the volume v is not influenced by any knowledge concerning the regionalization $z(x)$. In this case, exact mathematical results can be obtained, although the problem of the change of support remains extremely difficult.

Fundamentally, this problem is of a physical, not statistical nature, even if the language used here is probabilistic. Naturally, estimation problems also arise, but the influence of the support effect upon the recoverable reserves is generally much more important than any estimation error. The basic problem is to predict how the distributions, conditional or otherwise, assumed to be known for a point or quasi point support, will be altered under a change of support. This physical law that we have to infer involves the structure of the multivariate distributions in an essential way. As long as this basic problem is sidestepped, it is possible to dream about a "distribution free" geostatistics. See Journel (1983). If the change of support is taken into account as it must be in practice, speaking of a distribution free geostatistics sounds as unrealistic as, say, model free physics.

After an introductory example of a non-probabilistic nature, the definition of the diffusion type random functions and the expression for the variation δf of the p.d.f. are given in Section 2. Examples are presented in Section 3. In section 4, the predictions from the usual approximate models are compared with true distributions. Finally, the discontinuous case is examined, and a more general model is suggested.

1. AN INTRODUCTORY EXAMPLE

This first example is purely deterministic, not probabilistic. Its aim is to emphasize the physical nature of the change of support. For simplicity, we restrict ourselves to the one-dimensional case $x \in \mathbb{R}$, but the multidimensional case can be handled in the same way, as will be done in Section 3.3, but in a probabilistic framework.

We consider a function, or regionalized variable $z(x)$, $x \in \mathbb{R}$, which has continuous derivatives z', z'' and z'''. To eliminate any edge effect, this function $z(x)$ is assumed to be periodic, with a period L. Under these hypotheses, we may write :

$$(1\text{-}1) \qquad z(x+h) - z(x) = h\, z'(x) + \frac{h^2}{2}\, z''(x) + \frac{h^3}{6}\, R(x+h)$$

where the function R is bounded by a constant C independent of h.

Now, if the point x is replaced by a random variable $\underline{x}$ uniformly distributed on $(0,L)$, $z(\underline{x})$, $z'(\underline{x})$... become random variables. For simplicity, we assume that the distribution F of $z(\underline{x})$ has a p.d.f. $f(z)$. This p.d.f. $f(z)$ is defined by the relation

$$(1\text{-}2) \qquad \int \varphi(z)\, f(z)\, dz = \frac{1}{L} \int_0^L \varphi[z(x)]\, dx$$

which holds for any regular enough function φ.

The law of the change of support depends on the behavior of the function z in the neighborhood of the random point $\underline{x}$; so we must examine the variables $z(\underline{x}+h)$ with $|h| \le \varepsilon$ for a given $\varepsilon > 0$. Our tools will be the conditional drift and the conditional variogram.

1.1 The infinitesimal conditional drift.

From relation (1-1), the conditional drift is of the form :

$$E[z(x+\underline{h}) - z(x)/z(\underline{x}) = z] = b(z;h) + O(\varepsilon^2)$$

where the function b(z;h) is defined by the relation :

$$\int \varphi(z)\ b(z;h)\ f(z)\ dz = \frac{1}{L} \int_0^L [h\ z'(x) + \frac{h^2}{2}\ z''(x)]\ \varphi[z(x)]dx$$

which holds for any regular enough function φ.

But we have :

$$\int_0^L \varphi[z(x)]\ z'(x)\ dx = \int_{z(o)}^{z(L)} \varphi(z)\ dz = 0$$

because the function z(x) is periodic. In the same way, we find :

$$\int_0^L z''(x)\ \varphi[z(x)]\ dx = - \int_0^L [z'(x)]^2\ \varphi'[z(x)]\ dx$$

Thus, the <u>infinitesimal conditional drift</u> b(z;h) is of the form :

$$(1-3)\qquad \begin{cases} b(z;h) = h^2\ b(z) \\ b(z) = \frac{1}{2}\ E[z''(\underline{x})/z(\underline{x}) = z] \end{cases}$$

and the function b(z) is defined by :

$$(1-4)\qquad \int \varphi(z)\ b(z)\ f(z)\ dz = - \frac{1}{2L} \int_0^L [z'(x)]^2\ \varphi'[z(x)]\ dx$$

1.2 The infinitesimal conditional variogram.

In the same way, for $|h|$, $|h'| \leq \varepsilon$, the conditional covariance is of the form :

$$E[(z(\underline{x}+h) - z(\underline{x}))(z(\underline{x}+h')-z(\underline{x}))/z(\underline{x}) = z] = C(z;h,h') + O(\varepsilon^2)$$

and the infinitesimal conditional covariance $C(z;h,h')$ is :

$$(1\text{-}5) \qquad \begin{cases} C(z;h,h') = 2\ h\ h'\ a(z) \\ a(z) = \frac{1}{2}\ E[z'^2(\underline{x})/\ z(\underline{x}) = z] \end{cases}$$

with the function $a(z)$ defined by :

$$(1\text{-}6) \qquad \int \varphi(z)\ a(z)\ f(z)\ dz = \frac{1}{2L} \int_0^L [z'(x)]^2\ \varphi[z(x)]\ dx$$

If $h = h'$, we find :

$$\frac{1}{2}\ E[(z(\underline{x}+h) - z(x))^2/z(\underline{x}) = z] = h^2\ a(z) + O(\varepsilon^2)$$

so that the <u>infinitesimal conditional variogram</u> is :

$$\gamma(z;h) = h^2\ a(z)$$

so the first relation (1-5) may be rewritten in the form :

$$(1\text{-}7) \qquad C(z;h,h') = \gamma(z;h) + \gamma(z;h') - \gamma(z;h-h')$$

This means that our conditional process behaves like an intrinsic random function in the neighborhood of the conditioning point.

1.3 <u>A differential equation for the p.d.f.</u>

Now, by comparing (1-4) and (1-6), we find :

$$\int \varphi\ b\ f\ dz = - \int \varphi'\ a\ f\ dz$$

By integrating by parts, it follows :

$$\int \varphi \; b \; f \; dz = \int \varphi \frac{d}{dz} (af) \; dz$$

for any function φ. Thus, the functions a and b and the p.d.f. f(z) satisfy the following differential equation

$$\boxed{b \; f = \frac{d}{dz} (a \; f)} \qquad (1\text{-}8)$$

In other words, the p.d.f. is determined once the infinitesimal conditional drift and variogram are known, and we have :

$$f(z) = \frac{C}{a(z)} \; \exp\left(\int_0^z \frac{b(u)}{a(u)} \; du \right)$$

where C is a suitable normalizing constant.

1.4 Moments of order > 2.

The moments of order > 2 may be neglected, because it follows from (1-1) that we have for any $|h_i| \leq \varepsilon$, $i = 1, 2 \ldots k > 2$:

$$E \left[\prod_{i=1}^{k>2} \; | \; z(\underline{x}_i + h_i) - z(\underline{x}) | / z(\underline{x}) = z \right] = O(\varepsilon^2) \qquad (1\text{-}9)$$

1.5 The law for a small change of support

From these results, it is possible to find the expression for the variation δf of the p.d.f. under a small change of support. Let $\ell = \varepsilon$ be an infinitely small length, and consider the regionalized variable z_ℓ defined by :

$$z_\ell(x) = \frac{1}{\ell} \int_{-\frac{\ell}{2}}^{\frac{\ell}{2}} z(x+h) \; dh$$

From (1-1), we may write :

$$z_\ell(x) = z(x) + \frac{\ell^2}{24} z''(x) + O(\ell^2)$$

Then for any function φ which is bounded and has bounded derivatives φ' and φ'', for instance a complex exponential, we have :

$$\varphi[z_\ell(x)] = \varphi[z(x)] + \frac{\ell^2}{24} z''(x)\ \varphi'[z(x)] + O(\ell^2)$$

If f_ℓ is the p.d.f. of the variable $z_\ell(\underline{x})$, and $\delta f = f_\ell - f$ is the variation of the p.d.f., it follows :

$$\int \varphi(z)\ \delta f(z)\ dz = \frac{\ell^2}{24}\ \frac{1}{L} \int_0^L z''(x)\ \varphi'[z(x)]\ dx$$

$$= - \frac{\ell^2}{24}\ \frac{1}{L} \int_0^L [z'(x)]^2\ \varphi''[z(x)]\ dx$$

By comparing with (1-6) and integrating by parts, we conclude :

$$\int \varphi(z)\ \delta f(z)\ dz = - \frac{\ell^2}{12} \int \varphi''(z)\ a(z)\ f(z)\ dz$$

$$= - \frac{\ell^2}{12} \int \varphi(z)\ \frac{d^2}{dz^2}\ (af)\ dz$$

and finally, by also taking into account the differential equation (1-8) :

$$\text{(1-10)} \qquad \boxed{\delta f = - \frac{\ell^2}{12}\ \frac{d^2}{dz^2}\ (af) = - \frac{\ell^2}{12}\ \frac{d}{dz}\ (bf)}$$

In geostatistics, the <u>tonnage</u> $T(z)$ selected above a given cut-off grade z, and the corresponding conventional <u>benefit</u> function $B(z)$ (Matheron, 1983a), are defined by :

$$\begin{cases} T(z) = 1 - F(z) = \int_z^\infty f(u)\ du \\ B(z) = \int_z^\infty (u-z)\ f(u)\ du = \int_z^\infty T(u)\ du \end{cases}$$

By (1-10), the variations of these functions under a small change of support are :

(1-11) $$\delta T = \frac{\ell^2}{12}\, b\, f \quad ; \quad \delta B = -\frac{\ell^2}{12}\, a\, f$$

In other words, under a small change of support and for a given cut-off grade z, the variation of the tonnage is proportionnal to the conditional drift, and the loss (in dollars!) is proportional to the conditional variogram.

Another important geostatistical function is the function Q(T) which represents the quantity of recoverable metal for a given selected tonnage T. It is defined by

$$Q(T) = B[z(T)] + T\, z(T)$$

where z(T) is the inverse function of T(z). Thus, for a fixed T, the variation of Q(T) is :

$$\delta Q(T) = \delta B(z) + \delta z(T)\, \frac{\partial B}{\partial z} + T\, \delta z(T)$$

But we have $\frac{\partial B}{\partial z} = -\, T$, and finally :

(1-12) $$\delta Q(T) = \delta B(z) = -\frac{\ell^2}{12}\, a\, f$$ (with $z = z(T)$)

Under a small change of support and for a given selected tonnage T, the loss in metal is also proportional to the conditional variogram.

2. DIFFUSION TYPE RANDOM FUNCTIONS

These fundamental results, i.e. the differential equation (1-8) and the expression (1-10) for the law for a small change of support, were obtained under strong regularity hypotheses. We shall now generalize them inside a much wider framework, but this will require the use of probabilistic language. In a sense, our goal is a multidimensional and non-Markovian generalization of the Ito stochastic integral. Roughly speaking, it will be assumed that, in the neighborhood of a conditioning point, the conditional random function behaves like a Gaussian intrinsic random function.

The physical implications of this model are emphasized by the terminology : diffusion type random function. It may be said that this has implicitly worked as an heuristic model since the beginning, so that it might be responsible for the successes of Geostatistics up to now.

2.1 The Ito conditions.

We consider a stationary, but in general not multigaussian, random function $Z(x)$, $x \in \mathbb{R}^n$, and we want to study the behaviour of the conditional version of $Z(x)$ in the neighborhood of a given conditioning point x_o. Without loss of generality, this point x_o may be chosen as the origin of the coordinates, i.e. $x_o = 0$. More precisely, if <u>v is a small neighborhood</u> of $x_o = 0$ and $\gamma(h)$ is the variogram of the R.F. $Z(x)$, we put :

$$\varepsilon = \text{Sup } \{\gamma(x-y) \; ; \; x, y \in v\}$$

and the notation $O(\varepsilon)$ will denote a quantity such that $O(\varepsilon)/\varepsilon \to 0$ if $\varepsilon \to 0$.

The following conditions, which will be called Ito conditions, even though they are not in quite the standard form, are assumed to be satisfied for any $x, y, x_i \in v$:

$$(2\text{-}1)\qquad \begin{cases} E[Z_x - Z_o/Z_o = z] = b(z;x) + R_1 \\ E[(Z_x-Z_o)(Z_y-Z_o)/Z_o = z] = C(z;x,y) + R_2 \\ E[\prod_{i=1}^{k>2} |Z_{x_i} - Z_o|/Z_o = z] = R_3 \end{cases}$$

with remainders R_1, R_2, R_3 satisfying bounds of the form :

$$|R| \leq O(\varepsilon)\ H \quad ; \quad E[H^n] < \infty,\ n = 1,2,\ldots$$

We also assume that the random variable $b(Z_o;x)$ and $C(Z_o;x,y)$ have finite moments for all the required orders.

The function $b(z;x)$ is the infinitesimal conditional drift ; $C(z;x,y)$ is the infinitesimal conditional covariance, and the infinitesimal conditional variogram $\gamma(z;h)$ is defined by

$$\gamma(z;h) = \frac{1}{2}\, C(z;h,h) \qquad (h \in v)$$

2.2 A differential equation for the p.d.f.

Let φ be a bounded function with bounded derivatives φ', φ'' and φ'''. For instance, φ may be a complex exponential. Then, for any $x \in v$, we may write :

$$\varphi(Z_x) - \varphi(Z_o) = (Z_x-Z_o)\ \varphi'(Z_o) + \frac{1}{2}\,(Z_x-Z_o)^2\ \varphi''(Z_o) + \frac{(Z_x-Z_o)^3}{6}\ U$$

with a random variable U such that $|U| \leq \text{Sup} |\varphi'''(z)| < \infty$. It follows from the Ito conditions :

$$E[\varphi(Z_x) - \varphi(Z_o)/Z_o = z] = b(z;x)\ \varphi'(z) + \gamma(z,x)\ \varphi''(z) + R$$

with a remainder R satisfying a bound of the form (2-1'). By taking the expectation, we have :

$$E[\varphi(Z_x)] - E[\varphi(Z_o)] = E[b(Z;x)\ \varphi'(Z) + \gamma(Z;x)\ \varphi''(Z)] + O(\varepsilon)$$

But this expression is zero because of the stationarity. Thus we have :

$$\int b(z;x)\ \varphi'(z)\ f(z)\ dz = -\int \gamma(z;x)\ \varphi''(z)\ f(z)\ dz + O(\varepsilon)$$

$$= \int \varphi'(z) \frac{\partial}{\partial z} (\gamma(z;x)\ f(z))\ dz + O(\varepsilon)$$

The term $O(\varepsilon)$ may be dropped out, because the functions b and γ are defined up to $O(\varepsilon)$ type quantities only. This relation holds for any function φ, so that it follows :

(2-2) $$\boxed{b(z;x)\ f(z) = \frac{\partial}{\partial z} (\gamma(z;x)\ f(z))}$$

Thus, the p.d.f. is determined once the infinitesimal conditional drift and variogram are given. From another point of view, relation (2-2) also shows that the two functions b and γ do not behave independently. An equivalent form, which will be useful, is the following :

(2-2') $$E[b(Z;x)\ \varphi'(Z) + \gamma(Z;x)\ \varphi''(Z)] = 0$$

2.3 Structure of the conditional covariance

Let x and y be two points in v. Consider the random variables :

$$Z_\alpha = \alpha\ Z(x) + (1-\alpha)\ Z(y) = Z_o + \alpha(Z_x - Z_o) + (1-\alpha)(Z_y - Z_o)$$

Then, for any bounded function φ with bounded derivatives φ', φ'' and φ''', we have :

$$\varphi(Z_\alpha) = \varphi(Z_o) + (Z_\alpha - Z_o)\ \varphi'(Z_o) + \frac{1}{2}(Z_\alpha - Z_o)^2\ \varphi''(Z_o) + \frac{1}{6}(Z_\alpha - Z_o)^3\ U$$

with $|U| \leq \text{Sup}\ |\varphi'''| < \infty$. By the Ito conditions, it follows :

$$E[\varphi(Z_\alpha)/Z_o = z] = \varphi(z) + [\alpha\ b(z;x) + (1-\alpha)\ b(z;y)]\ \varphi'(z)$$
$$+ [\alpha^2\ \gamma(z;x) + (1-\alpha)^2\ \gamma(z;y) + \alpha(1-\alpha)\ C(z;x,y)]\ \varphi''(z) + R$$

and the remainder R satisfies a bound of the form (2-1'). Thus, by taking the expectation, we have :

$$E[\varphi(Z_\alpha)] = E[\varphi(Z_o)] + E[b(Z,y)\ \varphi'(Z) + \gamma(Z;y)\ \varphi''(Z)]$$
$$+ \alpha\ E[(b(Z,x) - b(Z,y))\ \varphi'(Z) + (C(Z,x,y) - 2\ \gamma(Z,y))\ \varphi''(Z)]$$
$$+ \alpha^2\ E[(\gamma(Z,x) + \gamma(Z,y) - C(Z;x,y))\ \varphi''(Z)] + O(\varepsilon)$$

From (2-2'), this may be rewritten :

$$E[\varphi(Z_\alpha)] - E[\varphi(Z_o)] = \alpha(1-\alpha)\ E[(C(Z;x,y) - \gamma(Z,x) - \gamma(Z,y))\ \varphi''(Z)]$$

Now, because of the stationarity, this expression does not change when x and y are replaced by x+h and y+h. In particular, if h = -y, we have $\gamma(Z;o) = C(Z;x-y,o) = 0$, and we find :

$$E[((C(Z;x,y) - \gamma(Z,x) - \gamma(Z,y))\varphi''(Z)] = -\ E[\gamma(Z;x-y)\ \varphi''(Z)]$$

for any regular enough function φ, and thus :

$$C(z;x,y) = \gamma(z,x) + \gamma(z,y) - \gamma(z;x-y) \tag{2-3}$$

This relation means that, in the neighborhood of the conditioning point $x_o = 0$, the conditional random function behaves like an intrinsic random function, except for the drift $b(z;x)$.

2.4 Variation of the p.d.f. under a small change of support.

Now, we consider the variable $Z_v = (1/v)\int_v Z(x)\,dx$, or, more generally, the variable :

$$Z_\mu = \int Z(x)\ \mu(dx) \qquad \left(\int \mu(dx) = 1\right)$$

where μ is a measure with its support $\subset v$ and such that $\int \mu(dx) = 1$. Then, for any regular enough function φ we find as above :

$$\varphi(Z_\mu) - \varphi(Z_o) = (Z_\mu - Z_o)\ \varphi'(Z_o) + \frac{1}{2}\,(Z_\mu - Z_o)^2\ \varphi''(Z_o) + \frac{1}{6}\,(Z_\mu - Z_o)^3\ U$$

$$|U| \leq \operatorname{Sup} |\varphi'''| < \infty$$

and, by the Ito conditions :

$$E[\varphi(Z_\mu)-\varphi(Z_o)] = E[\varphi'(Z)\int b(Z,x)\ \mu(dx) + \frac{1}{2}\ \varphi''(Z)\iint C(Z;x,y)\ \mu(dx)\ \mu(dy)] + O(\varepsilon)$$

But, from (2-3), we have

$$\frac{1}{2}\iint C(z;x,y)\ \mu(dx)\ \mu(dy) = \int \gamma(z,x)\ \mu(dx) - \frac{1}{2}\,\bar{\gamma}(z)$$

$$\bar{\gamma}(z) = \iint \gamma(z;x-y)\ \mu(dx)\ \mu(dy)$$

and from (2-2') :

$$E[\varphi'(Z)\int b(Z;x)\ \mu(dx) + \varphi''(Z)\int \gamma(z;x)\ \mu(dx)] = 0$$

so that we find :

(2-4) $$E[\varphi(Z_\mu)-\varphi(Z_o)] = -\frac{1}{2} E[\varphi''(Z)\ \overline{\gamma}(Z)] + O(\varepsilon)$$

Thus, the <u>variation</u> $\delta f = f_\mu - f$ of the p.d.f. satisfies the following relations :

$$\int \varphi(z)\ \delta f(z)\ dz = -\frac{1}{2}\int \varphi''\ \overline{\gamma}(z)\ f(z)\ dz = -\frac{1}{2}\int \varphi \frac{d^2}{dz^2}(\overline{\gamma}\ f)\ dz$$

for any regular enough function. It follows :

(2-4') $$\boxed{\delta f(z) = -\frac{1}{2}\frac{d^2}{dz^2}(\overline{\gamma}(z)\ f(z))}$$

From the differential equation (2-2), we find that the variation of the <u>tonnage</u> $T(z) = 1 - F(z)$ is given by :

$$\delta T(z) = \frac{1}{2}\overline{b}\ (z)\ f(z)$$

where $\overline{b}(z) = \iint b(z;x-y)\ \mu(dx)\ \mu(dy)$ is the average of the conditional drift. Concerning the function $B(z)$ and $Q(T)$, we find exactly as in Section 1 :

(2-5) $$\boxed{\delta B(z) = \delta Q(T) = -\frac{1}{2}\overline{\gamma}(z)\ f(z)}$$

<u>N.B.</u> From the relation $\int B(z)\ dz = \frac{1}{2} E(Z^2)$, (Matheron, 1983a), we always have :

$$\int \delta B(z)\ dz = \frac{1}{2}\delta\sigma^2 = -\frac{1}{2}\overline{\gamma} = -\frac{1}{2}\int \overline{\gamma}(z)\ f(z)\ dz$$

The relation (2-5) shows that this general rule also holds in the conditional sense, in the case of a diffusion type random function. It means that, for a given cut-off grade z, the loss due to a small change of support v is proportional to the average value in v of the conditional variogram.

2.5 <u>The case of the proportional effect.</u>

In the general case, the behavior of the conditional variogram $\gamma(z;h)$, and in particular its kind of anisotropy, may depend in a complex way on the conditioning value z. The simplest possible model is the case of the proportional effect, i.e. :

$$\gamma(z;h) = a(z)\ \gamma(h)$$

where $\bar{\gamma}(h)$ is the non conditional variogram. In this case, we also find

$$b(z;h) = b(z)\ \gamma(h)$$

and the functions a, b and f satisfy the differential equation :

(2-6)
$$\boxed{\frac{d}{dz}(a\ f) - b\ f = 0}$$

In particular, the p.d.f. is of the form :

$$f(z) = \frac{C}{a(z)}\ \exp\left(\int_0^z \frac{b(u)}{a(u)}\,du\right)$$

Consequently, the variations of the functions f(z), T(z), B(z) and Q(T) under a small change of support, are :

$$\begin{cases} \delta f(z) = -\frac{\overline{\gamma}}{2} \frac{d}{dz}(b\ f) = -\frac{\overline{\gamma}}{2} \frac{ab' - ba' + b^2}{a} f \\ \delta T(z) = \frac{\overline{\gamma}}{2}\ bf \\ \delta B(z) = \delta Q(T) = -\frac{\overline{\gamma}}{2}\ a\ f \end{cases}$$

If $u(z)$ is a regular enough function, we have :

(2-7) $$\delta\ E[u(Z)] = \frac{\overline{\gamma}}{2} E[b(Z)\ u'(Z)] = -\frac{\overline{\gamma}}{2} E[a(Z)\ u''(Z)]$$

and in particular the variation of the moment of order n is :

(2-7') $$\delta\ M_n = \frac{n}{2}\ \overline{\gamma}\ E[Z^{n-1}\ b(Z)] = -\frac{n(n-1)}{2}\ \overline{\gamma}\ E[Z^{n-2}\ a(Z)]$$

Note that we have :

(2-8) $$E[a(Z)] = 1$$

because, by definition, $\gamma(z;h) = a(z)\ \gamma(h)$ and $\gamma(h) = E[\gamma(Z;h)]$. Thus, in the particular case n = 2, (2-7') becomes :

$$\delta\ M_2 = \delta\sigma^2 = -\ \overline{\gamma}$$

as expected.

This very simple model of the proportional effect will probably prove to be very useful in the applications. But it must be kept in mind that more complex models are also possible.

3. EXAMPLES

It is not obvious a priori that stationary diffusion type random functions in $\mathbb{R}^n$ do exist. I shall give three examples, but I hope that many others do exist.

3.1 Diffusion Markov processes on the straight line

This first example is obvious, because it is the starting point of the whole theory. It is the stationary solution to the Ito stochastic differential equation :

$$d\ Z_t = b(Z_t)\ dt + \sqrt{a(Z_t)}\ d\xi_t$$

where ξ_t is a standard brownian motion, with $E(\xi_t) = 0$ and $Var(\xi_t) = 2\ t$. The Ito conditions are satisfied, and we are in the case of the proportional effect (see section 2-5). The only required condition is the existence of the stationary p.d.f. $f(z)$, i.e. :

$$\int \frac{1}{a(z)} \exp\left(\int_0^z \frac{a(u)}{b(u)}\ du\right) dz < \infty$$

The <u>infinitesimal generator</u> A of the Markov process Z_t is the operator :

$$A = a(z) \frac{d^2}{dz^2} + b(z) \frac{d}{dz} \tag{3-1}$$

For any regular enough function φ, the function

$$\varphi_t(z) = E[\varphi(Z_t)/Z_o = z]$$

is the unique solution of the <u>evolution equation</u>

$$\frac{d\ \varphi_t}{dt} = A\ \varphi_t \qquad (\varphi_o = \varphi)$$

so that we can write

$$\varphi_t = P_t\ \varphi = e^{At}\ \varphi$$

or explicitly :

$$\varphi_t(z) = \int P_t(z;dz') \; \varphi(z')$$

where $P_t(z;dz')$ is the transition probability of the Markov process Z_t. This means that the stationary bivariate distribution

$$f_t(z_o,z') \; dz_o \; dz' = f(z_o) \; P_t(z_o;dz')$$

of (Z_o,Z_t) is determined once the operator A, i.e. the two functions a(z) and b(z), are known.

Now, if Z(x) is a diffusion type random function in $\mathbb{R}^n$ with a proportional effect, we may also consider the Markov process Z_t on $\mathbb{R}$ defined by the operator (3-1), with the same functions a and b. The p.d.f. f(z) is the same for Z(x) and Z_t. But what happens to the bivariate distributions ? Are they of the same form, or more precisely, for any $h = x-y \in \mathbb{R}^n$, is it possible to find a value of t such that :

$$(3\text{-}2) \qquad f_h(z_x,z_y) = f_t(z_o,z_t) \qquad ?$$

The answer to this question (3-2) is very important for the practical applications, because any consistent non-linear estimation procedure requires the knowledge of at least the bivariate distributions (Z_x,Z_y) and (Z_x,Z_v).

The answer is in the affirmative in the multigaussian case, which will be our second example. It is doubtful in the case of the third example, and probably negative in the general case. For this reason, I suggest a "call for research" of the form : under what condition is the answer to the question (3-2) positive ?

3.2 The multigaussian case.

Let Y(x) be a stationary multigaussian random function. Without loss of generality, we may assume

$$E(Y(x)) = 0 \quad ; \quad Var(Y(x)) = 1$$

$$Cov(Y_x, Y_y) = \rho(x-y) = 1 - \gamma_o(x-y)$$

We assume that there is no nugget effect, i.e. the variogram γ_o is continuous. Then, the Ito conditions are satisfied with a proportional effect, and we have :

$$b_o(y;x) = - y\ \gamma_o(x) \quad ; \quad \gamma_o(y;h) = \gamma_o(h)$$

Now, we may define a new random function by putting :

$$Z(x) = \varphi(Y(x))$$

Then, if the transform φ is regular enough, Z(x) is also a diffusion type random function with a proportional effect. If we assume for instance that φ, φ', φ'' and φ''' satisfy an exponential bound of the form :

$$|\varphi| + |\varphi'| + |\varphi''| + |\varphi'''| \leq B \exp(C|y|)$$

it is easy to show that the Ito conditions are satisfied and that we have a proportional effect :

$$b(z;x) = b(z)\ \gamma_o(h) \quad ; \quad \gamma(z;h) = a(z)\ \gamma_o(h)$$

Note that these functions a and b are not exactly the same as in Section 2.5, because the variogram $\gamma(h)$ of Z(x) is proportional but not equal to $\gamma_o(h)$. If A_o is the generator of the standard Gaussian Markovian process Y_t, i.e.

$$A_o = \frac{d^2}{dy^2} - y \frac{d}{dy}$$

the function a(z) and b(z) are given by :

(3-3) $b(z) = E[A_o \, \varphi(Y)/\varphi(y) = z]$; $a(z) = E[\varphi'^2(Y)/\varphi(Y) = z]$

The appearance of the conditional expectation is due to the fact that generally the equation $\varphi(y) = z$ may have more than one solution. If the transform φ is a true anamorphosis, i.e. is strictly increasing, the conditional expectation can be dropped out, and we have :

(3-3') $$b(z) = A_o \, \varphi(y) \; ; \quad a(z) = \varphi'^2(y)$$

with y uniquely determined by $z = \varphi(y)$.

Note that the transformations (3-3) or (3-3') are general : if Y(x) is a diffusion type R.F. with a proportional effect defined by two functions $a_o(y)$ and $b_o(y)$, and if the transform φ is regular enough, the R.F. $Z(x) = \varphi(Y(x))$ is also of the diffusion type with a proportional effect, and the new functions a(z), b(z) are given by

$$b(z) = E[A_o \, \varphi(Y)/\varphi(Y) = z] \quad ; \quad a(z) = E[a_o(Y) \, \varphi'^2(Y)/\varphi(Y) = z]$$

$$\left(A_o = a_o(y) \frac{d^2}{dy^2} + b_o(y) \frac{d}{dy}\right)$$

3.3 The regular case

Now, we suppose the stationary R.F. Z(x) to be three times differentiable in quadratic mean. More precisely, we assume that :

$$Z(x) = Z_o + x^i \partial_i Z_o + \frac{1}{2} x^i x^j \partial_{ij} Z_o + O(|x|) R$$

with $|R| \leq H$ for a random variable H independent of the point x and such that $E[H^n] < \infty$ for any n. We also assume that the partial derivatives $\partial_i Z_o$, $\partial_{ij} Z_o$ and Z_o itself have finite moments of all required ordres.

Then, it is easy to show that the Ito conditions are satisfied with :

$$\begin{cases} b(z;x) = x^i x^j b_{ij}(z) \quad ; \quad b_{ij}(z) = \frac{1}{2} E[\partial_{ij} Z_o / Z_o = z] \\ \gamma(z;h) = h^i h^j a_{ij}(z) \quad ; \quad a_{ij}(z) = \frac{1}{2} E[\partial_i Z_o \partial_j Z_o / Z_o = z] \end{cases}$$

In the case of a proportional effect, there is a constant positive definite matrix K_{ij} such that

$$a_{ij}(z) = a(z) K_{ij} \quad ; \quad b_{ij}(z) = b(z) K_{ij}$$

But there is no reason for these relations to be satisfied in the general case.

4. APPROXIMATE MODELS OF CHANGE OF SUPPORT

The preceding results, and in particular the formula (2-4) can be used only in the case of a very small change of support. But in practice the variance reduction is often greater than 50%. In the case of a relatively large change of support, two kinds of approximate models are available :

- the affine correction (Journel and Huijbregts, 1978)
- the isofactorial models (Matheron, 1983b), especially

popular in the very particular case called discretized Gaussian model (Matheron, 1975b and Journel and Huijbregts, 1978). A more general presentation is given in Matheron(1983b) in the discrete case.

In this section, we shall compare the predictions of these models with the true distributions, in the case of a diffusion type random function with proportional effect. The comparison will be made for the first order approximation in the general case, and for the second order in the multigaussian case. Moreover, we shall give actual values in a few particular examples. The general conclusions will be as follows :

~ Isofactorial model predictions are exact for the first order approximation and, in the multigaussian case, are almost correct for the second order approximation.

~ Affine correction is false for the first order onwards, except if $z - m$ is a factor, i.e. an eigen function for the operator $A = a(z) \frac{d^2}{dz^2} + b(z) \frac{d}{dz}$.

4.1 The affine correction (A.C.)

In this very simple model, it is assumed that $(Z_v-m)/\sigma_v$ and $(Z_o-m)/\sigma$ have the same distribution. Thus, by putting

$$\varepsilon = 1 - \frac{\sigma_v}{\sigma} = \frac{\bar{\gamma}}{2\sigma^2} + O(\bar{\gamma})$$

Z_v and $Z_o - \varepsilon\,(Z_o-m)$ have the same distribution, and we find, for any regular enough function $u(z)$:

$$E_{Ac}[u(Z_v)] = E[u(Z_o)] - \frac{\bar{\gamma}}{2\sigma^2}\, E[(Z_o-m)\, u'(Z_o)] + O(\bar{\gamma})$$

By comparison with (4-1), we see that this expression is correct only if we have

$$E[b(Z)\ u'(Z)] = -\frac{1}{\sigma^2} E[(Z-m)\ u'(Z)]$$

for any function u, i.e. :

$$b(z) = -\frac{Z_o-m}{\sigma^2}$$

In terms of the differential operator

$$A = a(z)\frac{d^2}{dz^2} + b(z)\frac{d}{dz}$$

associated with the diffusion type R.F. $Z(x)$, this condition may be written as :

$$A(z-m) = -\lambda(z-m)$$

Hence the conclusion : the affine correction is false for the first order, except if (z-m) is an eigen function of the operator A.

In the multigaussian case, of Section 3.2, the conclusion is as follows :

If φ is a true anamorphosis, the affine correction is exact for the first order if and only if

$$A_o(\varphi-m) = -\lambda(\varphi-m)$$

i.e. if φ-m is an Hermite's polynomial : but Hermite's polynomials of degree $n \geq 2$ are not monotonic, so that in fact the affine correction is false for any non-linear anamorphosis.

Now, if φ is a regular transform, but not strictly monotonic, , the affine correction is exact for the first order if φ-m is an Hermite's polynomial, or more generally, if

$$E[A_o \varphi(Y)/\varphi(Y) = z] = -\lambda(z-m)$$

This condition is satisfied for instance, if $\varphi(y) = y^2$, and also if $\varphi(y) = y^2 \, 1_{y>o}$.

4.2 The Isofactorial model

The starting point of this model is the Cartier condition :

$$E[Z_{\underline{x}}/Z_v] = Z_v$$

where the random point $\underline{x}$ is uniformly distributed inside v (Matheron, 1983a et b). Moreover, the bivariate distribution of $(Z_{\underline{x}}, Z_v)$ is assumed to be the same as if we had

$$Z_v = \varphi_v(Z_{t_o}) \quad ; \quad Z_{\underline{x}} = Z_{t_o+\tau}$$

where Z_t is the diffusion Markov process defined by the generator

$$A = a(z) \frac{d^2}{dz^2} + b(z) \frac{d}{dz}$$

From Cartier's condition, this comes back to putting

$$Z_v = \varphi_v(Z) \quad ; \quad \varphi_v = P_t \, z = e^{At} \, z$$

so that, for the first order approximation, we find

$$\varphi_v(z) = z + t\,A\,z + O(t) = z + t\,b(z) + O(t)$$

The parameter t must be chosen so that the variance has the correct value σ_v^2 . For the first order approximation, we have :

$$E[\varphi_v^2(Z)] = E(Z^2) + 2\ t\ E(Z\ A(Z)) + O(t)$$

But, from the differential equation (2-6), we find

$$E[u(Z)\ A\ v(Z)] = -\ E[a(Z)\ u'(Z)\ v'(Z)]$$

and thus $E(Z\ A\ Z) = -\ E(a(Z)) = -\ 1$ from (2-8). It follows that the variance has the correct value if

$$t = \frac{\bar{\gamma}}{2}$$

and finally the model is :

$$\varphi_v(z) = z + \frac{\bar{\gamma}}{2}\ b(z) + O(\bar{\gamma})$$

For any regular enough function u, it follows :

$$E[u(\varphi_v(Z))] = E[u(Z)] + \frac{\bar{\gamma}}{2}\ E[b(Z)\ u'(Z)] + O(\bar{\gamma})$$

By comparison with (2-7), this is the correct value, and we conclude :

The isofactorial model is exact for the first order approximation.

4.3 The second order approximation

In the multigaussian case $Z(x) = \varphi(Y(x))$ of Section 3.2, it is possible to obtain the second order approximation formula for the variation δf of the p.d.f. $f(z)$. After routine calculations, we find for any regular enough function $u(z)$:

$$\begin{cases} E[u(Z_v)] = E[u(Z_o)] - \frac{\overline{\gamma}_o}{2} \; E[u''(Z_o)\; \varphi'^2(Y_o)] \\ \qquad + \frac{1}{2}\left[\frac{u''''.(\varphi')^4}{4}(\overline{\gamma}_o)^2 + u'''.(\varphi')^2\; \overline{\gamma_o\gamma_o'} + \frac{1}{2}\, u''.(\varphi'')^2\; \overline{\gamma_o^2}\right] \\ \qquad + 0(\overline{\gamma}^2) \end{cases}$$

In particular, if $u(z) = z^n$, the order n moment $M_n(v)$ is given by :

$$\begin{cases} M_n(v) = M_n(o) - \frac{n(n-1)}{2}\; \overline{\gamma}_o\; E[Z^{n-2}\; \varphi'^2(Y)] \\ \quad + \frac{1}{2}\, E\left[\frac{n(n-1)(n-2)(n-3)}{4}(\overline{\gamma}_o)^2\; Z^{n-4}\; \varphi'^4 + n(n-1)(n-2)\; \overline{\gamma_o\gamma_o'}\; Z^{n-3}\varphi'^2 \right. \\ \qquad \left. + \frac{n(n-1)}{2}\; Z^{n-2}\; \varphi''^2\; \overline{\gamma_o^2}\right] + 0(\overline{\gamma}^2) \end{cases}$$

Note the occurence of the term

$$\overline{\gamma_o\gamma_o'} = \frac{1}{v^3} \iiint \gamma_o(x-y)\; \gamma_o(x-y')\; dx\; dy\; dy'$$

For the second order approximation, we find after some calculations that the prediction of the isofactorial model is correct, except that this term $\overline{\gamma_o\gamma_o'}$ is replaced by $(\overline{\gamma}_o)^2$. From a numerical point of view, the difference is very small. For instance, in the one-dimensional case, if $\gamma_o(h) = |h| + 0(|h|^2)$, we find :

$$\overline{\gamma_o\gamma_o'} = \frac{7}{60}\, h^2 + 0(h^2) = 0{,}11\ 667\ h^2 + 0(h^2)$$

$$(\overline{\gamma}_o)^2 = \frac{1}{9}\, h^2 + 0(h^2) = 0{,}11\ 111\ h^2 + 0(h^2)$$

Hence the conclusion : in the multigaussian case, the isofactorial model is almost correct for the second order approximation.

4.4 The lognormal case

In this model, we have

$$Z(x) = m\, e^{sY(x) - \frac{s^2}{2}}$$

where Y(x) is a standardized stationary multigaussian random function, as in Section 3.2. In this case, the prediction of the isofactorial model is another lognormal distribution with the same mean m, i.e.

$$Z_v = m\, e^{s_v Y - s_v^2/2}$$

and the parameter s_v is determined so that the variance has the correct value σ_v^2 .

But in this case the true values of the order n moment $M_n(v)$ are

$$M_n(v) = E[Z_v^n] = \frac{1}{v^n} \int_{v^n} E[Z_{x_1} Z_{x_2} \ldots Z_{x_n}]\, dx_1 dx_2 \ldots dx_n$$

$$= \frac{1}{v^n} M_n(o) \int_{v^n} e^{-s^2 \sum_{i<j} \gamma_o(x_i - y_j)}\, dx_1\, dx_2 \ldots dx_n$$

so that numerical calculations are always possible, at least for n = 3 and 4.

In the one-dimensional case, if the variogram of Y(x) is

$$\gamma_o(h) = |h| \quad \text{if} \quad |h| \leq 1 \text{ , } 1 \text{ otherwise}$$

explicit formulae were obtained, after fairly tedious calculations. Figures 1, 2, 3 and 4 show the results in the case m = s = 1, for the variable

$$Z_v = \frac{1}{t} \int_0^t Z_x\, dx$$

Figure 1 shows the relative order 3 moment $M_3(t)/M_3(o)$. At this scale, true values and isofactorial predictions cannot be distinguished. It is surprising that the order two approximation formula is better than affine correction prediction.

Figure 2 - A much clearer discrimination is obtained by using the order 3 relative cumulant

$$\frac{\chi_3(t)}{\sigma^3(t)} = E\left[\left(\frac{Z_v-m}{\sigma_v}\right)^3\right]$$

The true values are decreasing because of the convergence towards normality. The isofactorial predictions are extremely good if $t \leq 1$. However in the affine correction model, χ_3/σ^3 remains constant, because $(Z_v-m)/\sigma_v$ and $(Z-m)/\sigma$ have the same distribution.

Figure 3 shows the order 4 relative cumulant χ_4/σ^4 : the conclusions are the same.

Figure 4 shows (on a logarithmic scale) the long run behavior of the relative cumulant χ_3/σ^3. From the general theory (convergence towards normality) we know that the true values are $\sim 1/\sqrt{t}$ for large enough t. The isofactorial prediction is also $\sim 1/\sqrt{t}$ (this is a general result). If t is ≥ 5 or 10, the ratio isofactorial prediction/true value remains constant $\simeq 0.78$, while the affine correction prediction remains constant !

N.B. In the multigaussian case, we saw that the isofactorial model is correct for a small change of support. It turns out that it is also correct for a very large change of support, because the convergence of $(\varphi_v(Y)-m)/\sigma_v$ towards a normal variable $N(0,1)$ automatically holds. This is the reason why the isofactorial model appears really good for any change of support in this case.

4.5 The gamma diffusion process

If $Y(t)$ is the Markov process defined by the generator

$$A = a(y) \frac{d^2}{dy^2} - b(y) \frac{d}{dy}$$

put :

$$H_t(\lambda;y) = E\left[e^{-\lambda \int_0^t \varphi(Y_\tau) d\tau} / Y_o = y \right]$$

This is the conditional Laplace transform of

$$t\, Z_v = \int_o^t \varphi(Y_\tau)\, d\tau$$

and the non conditional Laplace transform is :

$$\Phi_t(\lambda) = E\left[e^{-\lambda \int_0^t \varphi(Y_\tau) d\tau} \right] = E[H_t(\lambda, Y)]$$

In fact, $H_t(y;\lambda)$ is the solution to the equation :

$$\frac{\partial H_t}{\partial t} = A\, H_t - \lambda\, \varphi(y)\, H_t \qquad (H_o = 1)$$

We shall consider the diffusion process defined by

$$A = y \frac{d^2}{dy^2} + (\alpha - y) \frac{d}{dy}$$

This is the gamma process : the p.d.f. is the gamma (α) p.d.f.

~ As a first example, we choose $\varphi(y) = y$ itself, and the variable of interest is :

$$Y_v = \frac{1}{t} \int_o^t Y(\tau)\, d\tau$$

where $Y(t)$ is the gamma process.

By solving the equation

$$\frac{\partial H_t}{\partial t} = y \frac{\partial^2 H_t}{\partial^2 y} + (\alpha - y) \frac{\partial H_t}{\partial y} - \lambda y H_t \qquad (H_o = 1)$$

we find

$$(4\text{-}1) \qquad \Phi_t(\lambda) = E\left[e^{-\lambda Y_v t}\right] = \left[\frac{R\, e^{t/2}}{(R \operatorname{ch} \frac{Rt}{4} + \operatorname{Sh} \frac{Rt}{4})(\operatorname{ch} \frac{Rt}{4} + R \operatorname{Sh} \frac{Rt}{4})}\right]^{\alpha}$$

$$R = \sqrt{1+4\lambda})$$

This is the <u>rigorous solution</u>.

~ In this particular case, y-m is a <u>factor</u> for the gamma process, so that <u>the isofactorial and A.C. models coincide</u> and give the same prediction :

$$\Phi_t^*(\lambda) = e^{-\lambda\alpha t(1-\rho)} \frac{1}{[1+\lambda \rho t]^{\alpha}}$$

Numerically, Φ_t and Φ_t^* are very similar.

<u>Figure 5</u> shows the relative third cumulant $\chi_3(t)/\sigma^3(t)$ for this.

At the beginning, the true values decrease very slowly because y-m is a factor.

The common isofactorial and A.C. prediction is χ_3/σ^3 = constant. It is good if $t \leq 2$ or 2.5, or $\sigma^2(t) > \frac{1}{2}\sigma^2(o)$.

In the same figure we give the values of $\chi_3(t)/\sigma^3(t)$ in the case of the Ambarzumian process. It is as follows :

(exponential decrease and, from time to time, according to a Poisson process, a random exponentially distributed jump). The Ambarzumian process :

- ~ is Markov
- ~ has the same stationary gamma (α) distribution
- ~ has the same exponential covariance $e^{-|h|}$.

But y-m is not a factor for the Ambarzumian process. This is the reason why the relative cumulant decreases much faster.

Figure 6 : the same process in the long run : the common prediction of A.C. and isofactorial models becomes very poor after t = 5. The reason is the convergence of the true variable Y_v towards normality.

The true Laplace transform of Y_v, given by (3-9), has a Weierstrass type expansion as an infinite product :

$$\Phi_t(\lambda) = \prod_{n=1}^{\infty} \left(\frac{b_n}{b_n+\lambda} \right)^{\alpha}$$

i.e. : Y_v is an infinite sum of independent R.V. The eigen values b_n are known. In the case

$$\alpha = 1 \quad \text{i.e.} \quad g(y) = e^{-y}$$

we obtain an expansion of the form :

$$\Phi_t(\lambda) = \Sigma \frac{B_n}{b_n+\lambda}$$

so that an explicit inversion is possible. Numerical values are given in Matheron (1983b): the agreement between the true distribution and the isofactorial prediction is extraordinarily good.

4.6 The square Y_t^2 of the gamma process.

With the same gamma diffusion process Y_t , we now choose :

$$Z_t = \frac{Y_t^2}{m_2} \quad ; \quad Z_v = \frac{1}{t}\int_0^t \frac{Y_\tau^2}{m_2}\, d\tau$$

Here, $\varphi(y) = y^2/m_2$ is not a factor, so that we shall have a clear discrimination between the isofactorial and A.C. models.

Figure 7 - By using Laguerre polynomials, it is possible (although fairly tedious!) to calculate the true values of the third moment. This shows that

~ A.C. is very poor

~ The isofactorial model is very good for $\sigma_t^2 \geq \sigma_o^2/2$, but it is not correct if t is large : in this model, the variable $\varphi_v(Y)$ converges towards the first factor, which is gamma and not normal as it should.

5. A MORE GENERAL MODEL

In the discontinuous cases, the preceding models do not work. Let us give two simple examples.

First example. In the multigaussian case, if the transform φ is not continuous, our approximation formulae are no longer valid. For instance, if φ is an indicator function, i.e.

$$Z(x) = 1_{Y(x) \geq a}$$

we find

$$M_2(v) = 1 - G(a) - \frac{e^{-a^2/2}}{\sqrt{\pi}} \sqrt{\overline{\gamma}} + 0(\sqrt{\overline{\gamma}})$$

$$M_3(v) = 1 - G(a) - \frac{3}{2}\, e^{-a^2/2} \sqrt{\overline{\gamma}} + 0(\sqrt{\overline{\gamma}})$$

where G is the standard normal c.d.f. Note the occurence of the square root of the variogram, which is characteristic of the indicator function.

If Y(t) is the standard Gaussian Markovian process,,it is possible to calculate the general moment $M_n(t)$ of the variable

$$Z_v = \frac{1}{t} \int_0^t 1_{Y(\tau) > a}\, d\tau$$

After some fairly difficult calculations, we find

$$M_n(t) = 1 - G(o) - \frac{\sqrt{t}}{\pi}\, e^{-a^2/2} \left[\frac{4n}{1+2n} - 2^{n+1} \frac{(n!)^2}{(2n+1)!} \right] + 0(\sqrt{t})$$

This is compatible with neither the isofactorial nor the affine correction model.

Second example: Feller type Markov processes. These processes are jump processes, defined by an infinitesimal generator A of the form :

$$(A\,\varphi)(y) = -\,C(y)\,\varphi(y) + C(y) \int \Pi_y(du)\,\varphi(u)$$

where $\Pi_y(du)$ is a transition probability, and $C(y) \geq 0$.

A Feller type process is discontinuous, nevertheless it is possible to find the first order approximation for the variable

$$Z_v = \frac{1}{t} \int_0^t \varphi(Y_\tau)\, d\tau$$

For, if t is very small :

~ either Y_τ remains constant on $(0,t)$, with a probability $1 - t\,C(y) + 0(t)$

~ or it has only one jump, with a probability $t\,C(y) + 0(t)$

~ the probability of having more than one jump is $0(t)$.

Now this (eventual) unique jump point is uniformly distributed on $(0,t)$, so that, neglecting events of probability $0(t)$ we have :

$$\boxed{Z_V = U\,Z_o + (1-U)\,Z_t}$$

with U uniformly distributed on (0,1) and independent of the process. Clearly, this relation also holds if there is no jump, because in this case $Z_o = Z_t$.

This very simple result leads to :

$$\boxed{M_n(V) = M_n(o) + \frac{t}{n+1}\sum_{k=1}^{n-1} E[\varphi^k \, A \, \varphi^{n-k}] + 0(t)}$$

It turns out that this formula also holds in the case of the Ito type processes, and also for various other types of processes.

This suggests a new model, more general than the isofactorial model.

A new model : the uniform mixture.

If the bivariate distributions (Z_x, Z_y) are of the same type $g_\rho(z,z')$ except for the value of (say) the correlation coefficient ρ, put :

$$Z_V = U\, Z_o + (1-U)\, Z_1$$

~ U is uniformly distributed on (0,1) and independent of Z_o, Z_1

~ the distribution of (Z_o, Z_1) if $g_\rho(z_o, z_1)$ and ρ is chosen so that the variance has the correct value :

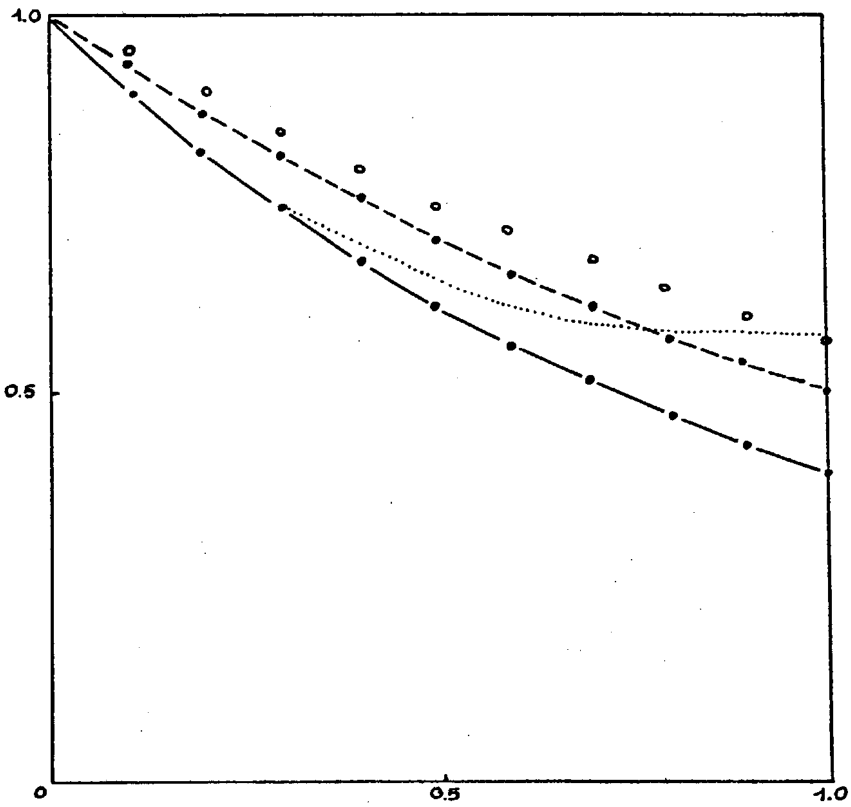
$$\sigma_V^2 = \frac{2+\rho}{3}\,\sigma^2$$

~ In practice, this model is good only if $\rho \geq 0$. Hence the limitation

$$\sigma_V^2 \geq \frac{2}{3}\,\sigma^2$$

If so, this uniform mixture model is practically equivalent to the isofactorial model, in the case of a diffusion type R.F. (and identical to it for the first order approximation). But it may be applied in other cases, in particular in discontinuous cases. New estimation techniques could be based upon it.

Figure 1



Lognormal Process 1-D - $M_3(t)/M_3(o)$

$s = 1$; $\sigma^2/m^2 = 1,71828..$; $\rho(h) = (1-|h|)_+$

—— · —— : true values of $M_3(t)/M_3(o)$, and isofactorial predictions : they cannot be distinguished at this scale !

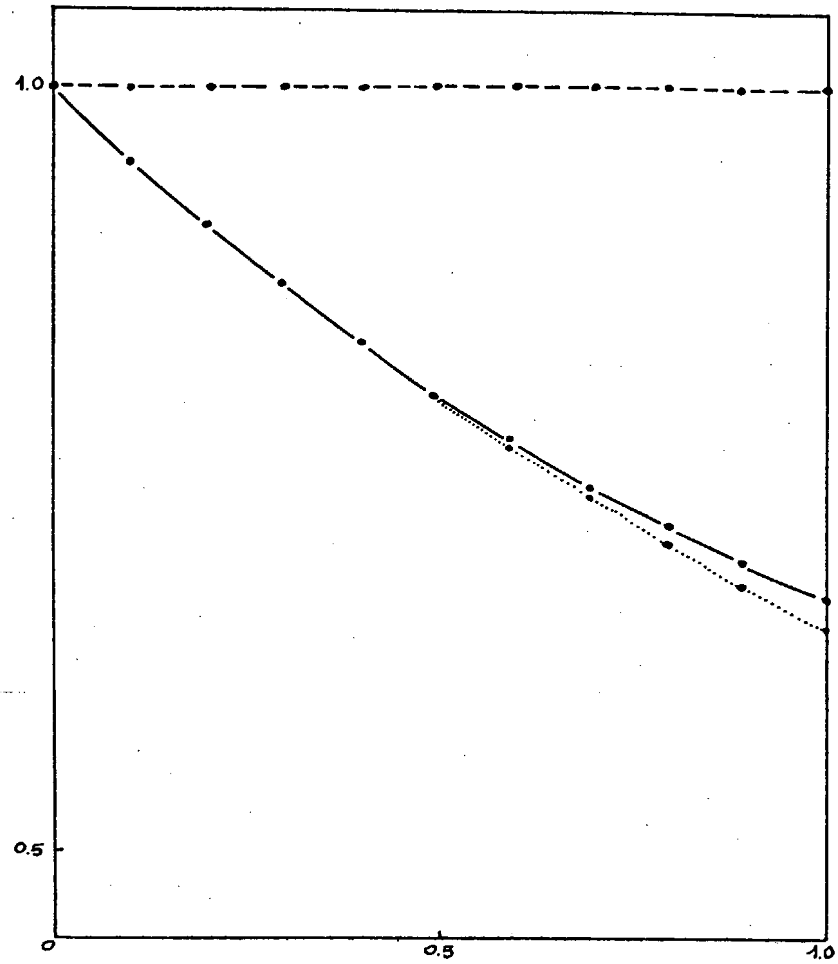
· — — · — — · : A.C. prediction

········· : Order two approximation : it is better than the A.C. prediction

● ● ● : The relative variance $\sigma^2(t)/\sigma^2(o)$

Figure 2

Lognormal Process



Order 3 relative cumulant $\chi_3(t)$

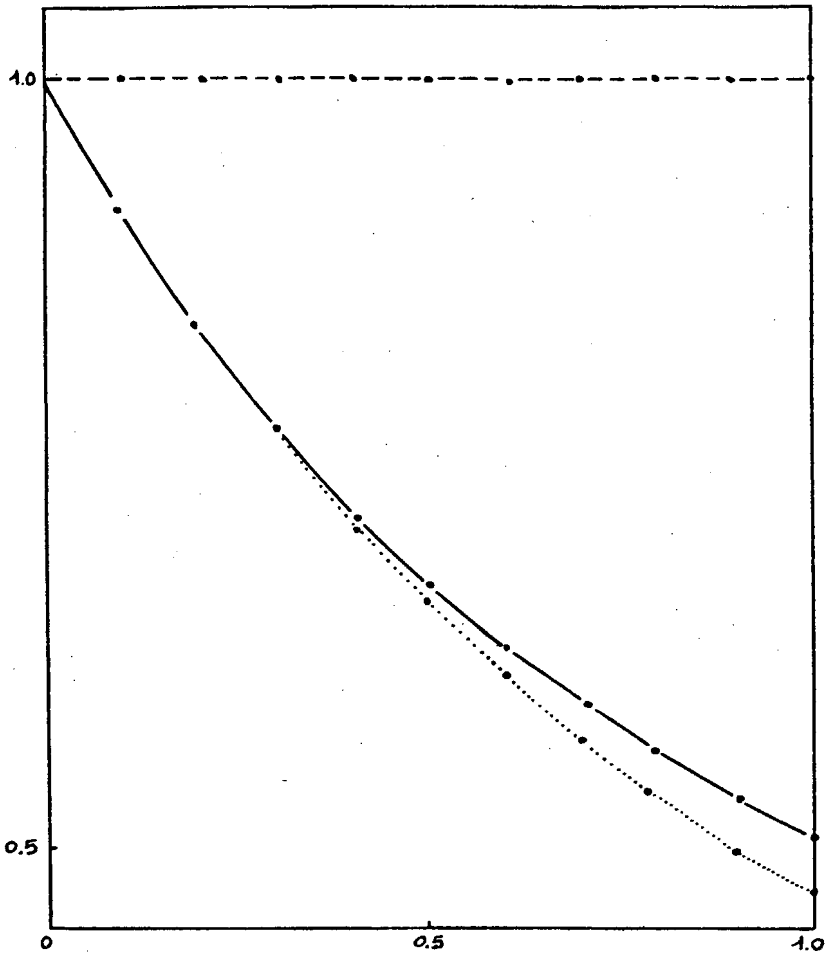
o ——— o ——— o : True values

o . . . . . o . . . . . o : Isofactorial predictions

o — — o — — o : A.C. predictions (remain constant !)

Figure 3

Lognormal Process 1-D



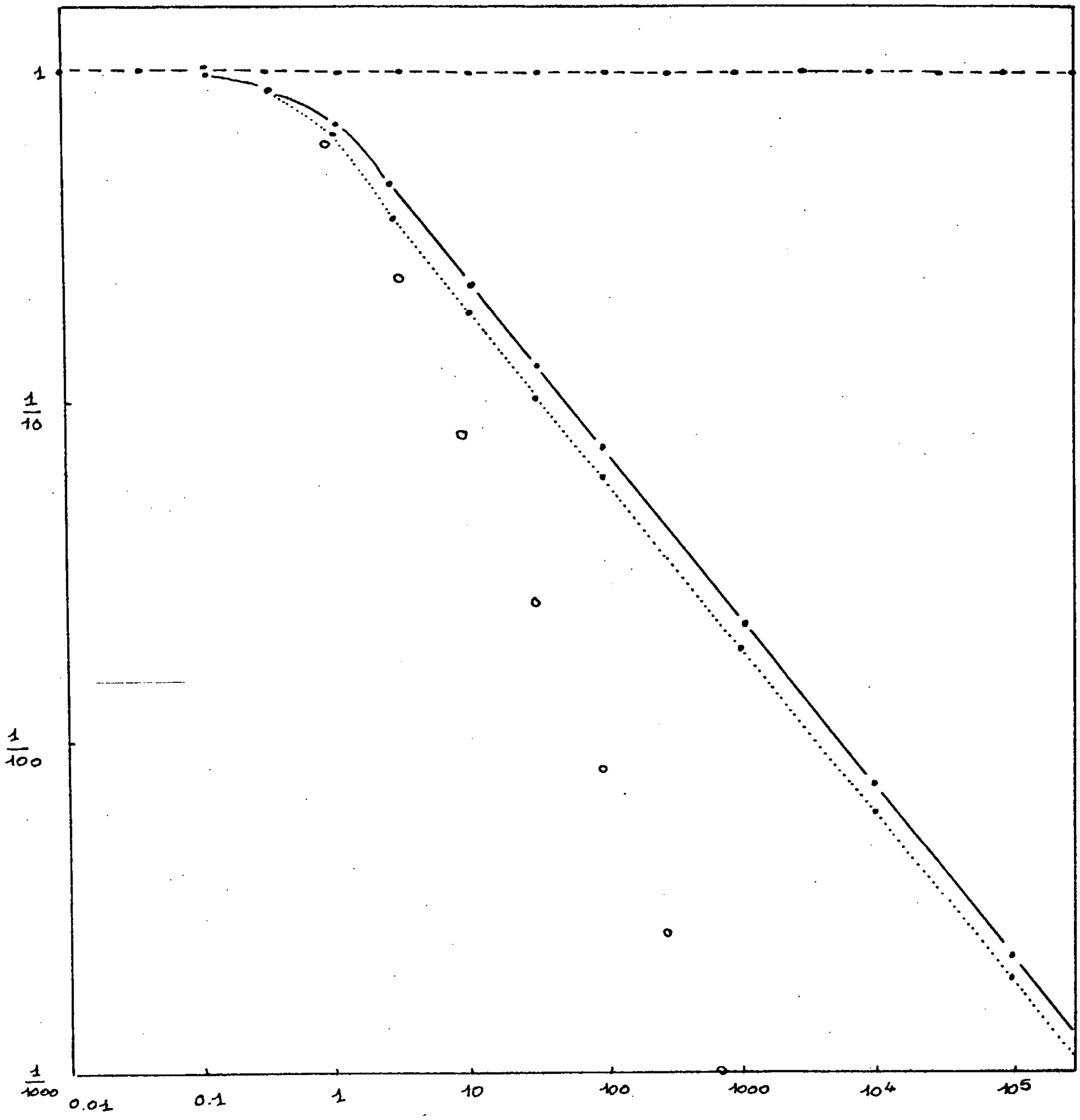
Order 4 relative cumulant $\chi_4(t)/\sigma^4(t)$

•——•——• : True values

•....•....• : Isofactorial predictions

•– –•– –• : A.C. predictions

Figure 4



The lognormal process in the long run

•——•——• : True values of $\chi_3(t)/\sigma^3(t)$: ~ $1/\sqrt{t}$ if $t \geq 10$

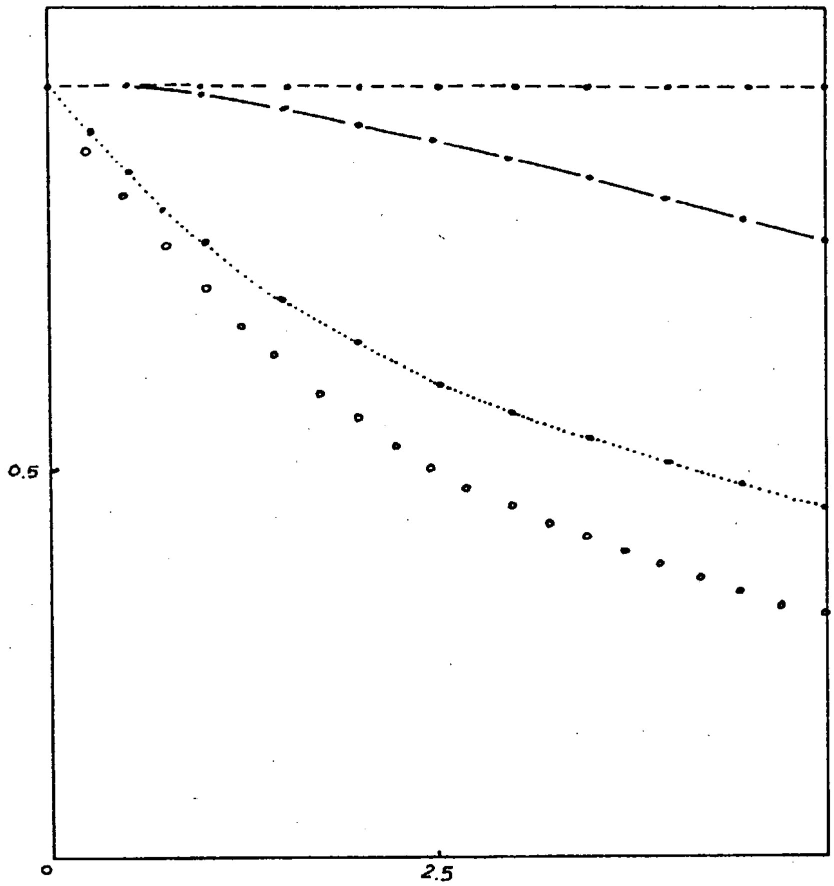
•....•....• : Isofactorial predictions ~ $1/\sqrt{t}$ if $t \geq 10$ and isofactorial/true value ≃ 0.78

•— —•— —• : Affine correction : remains constant !

• • • : Relative variance $\sigma^2(t)/\sigma^2(o)$ ~ $1/t$ if $t \to \infty$

Figure 5

The gamma Process Y_t



Order 3 relative cumulant $\chi_3(t)/\sigma^3(t)$ of the variable

$$Y_V = \frac{1}{V}\int_0^t Y(\tau)\, d\tau$$

•——•——• : True values

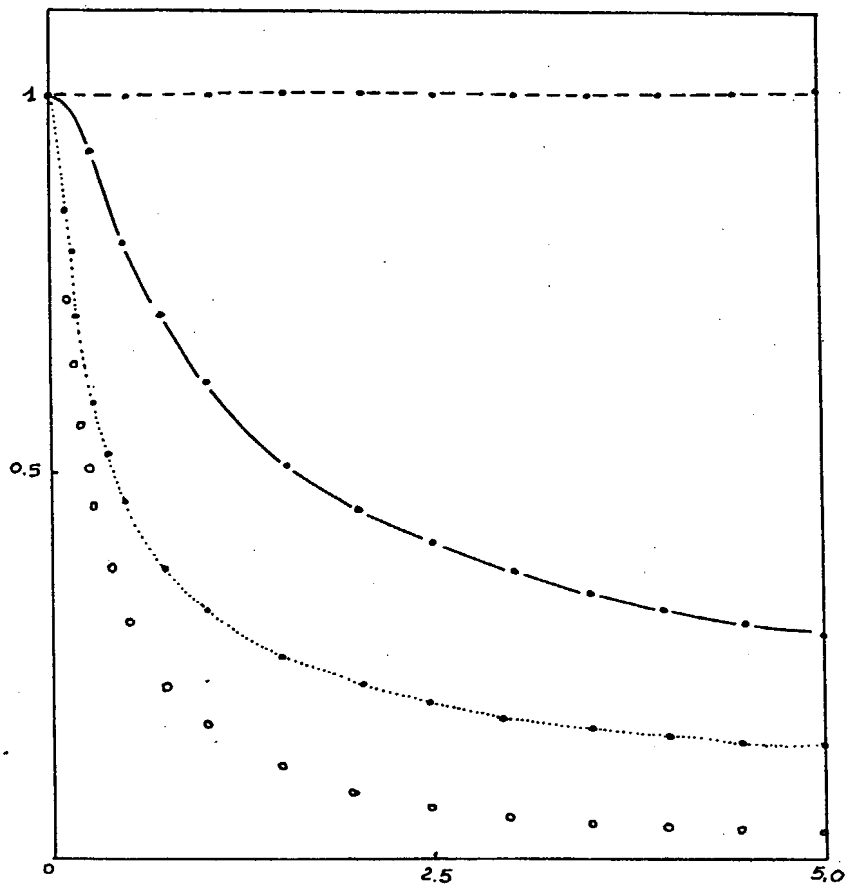
•— —•— —• : Common prediction of the isofactorial and A.C. models

• • • : The relative variance

•....•.....• : χ_3/σ^3 in the case of the Ambarzumian process (it is Markov, has the same gamma distribution and the same covariance $e^{-|h|}$: but y-m is not a factor for the Ambarzumian process).

Figure 6

The gamma Process Y_t in the long run



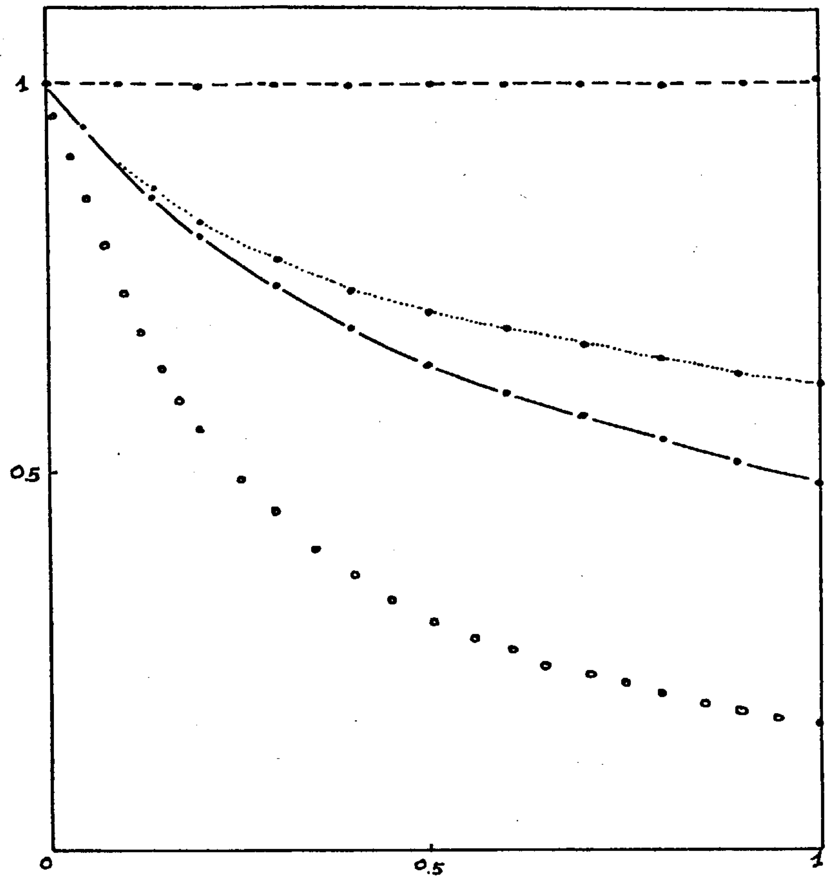
Order 3 relative cumulant $\chi_3(t)/\sigma^3(t)$ of the variable

$$Y_V = \frac{1}{t}\int_0^t Y(\tau)\, d\tau$$

- ○ —— ○ —— ○ : True values
- ○ — — ○ ——○ : Common prediction of isofactorial and A.C. models
- ●   ●   ● : The relative variance $\sigma^2(t)/\sigma^2(o)$
- ● . . . . ● . . . . ● : χ_3/σ^3 in the case of the Ambarzumian process.

Figure 7

The gamma Diffusion Process Y_t $\alpha = 5$



Relative 3rd cumulant of $Z_V = \frac{1}{t}\int_0^t \frac{Y_\tau^2}{m_2}\, d\tau$

•——•——•: True values

•.....•.....•: Isofactorial predictions

•— — •— — •: A.C. predictions (remain constant)

• • • • • • • •: The relative variance $\sigma^2(t)/\sigma^2(o)$

BIBLIOGRAPHY

1. JOURNEL, A.G. : Non-parametric estimation of spatial distributions. J. of Internat. Ass. for Mathematical Geology, Vol. 15, n° 3, pp. 445-468, June 1983.

2. MATHERON, G. : The selectivity of the distributions, and the second principle of geostatistics. (Internal Report N-800, C.G.M.M., March 1983. (1983a)

3. JOURNEL, A.G., and HUIJBREGTS, Ch. : "Mining Geostatistics", Academic Press, London, 1978.

4. MATHERON, G. : Isofactorial models and change of support. To appear in the Proceedings, 2nd NATO ASI "Geostat Tahoe 83", South Lake Tahoe, 4-17 Sept. 1983. (1983b).

5. MATHERON, G. : A simple substitute for conditional expectation : the disjunctive kriging. Proceedings, NATO ASI, "Advanced Geostatistics in the Mining Industry", Rome, Oct. 1975, pp. 221-326. (1975a).

6. MATHERON, G. : Forecasting block grade distributions : the transfer functions. Proceedings, NATO ASI "Advanced Geostatistics in the Mining Industry, Rome, October 1975, pp. 237-251. (1975b).