Mathematical Geology, Vol. 18, No. 8, 1986

Disjunctive Kriging Revisited: Part I[1]

M. Armstrong[2] and G. Matheron[2]

Difficulties in applying disjunctive kriging (D.K.) with an anamorphosis to a normal distribution have led to an interest in D.K. based on other distributions. After reviewing Gaussian D.K., this paper reviews other types of D.K. based on other infinitely divisible distributions (gamma, Poisson, and negative binomial).

KEY WORDS: nonlinear geostatistics, disjunctive kriging, gamma distribution, Poisson distribution, negative binomial distribution, isofactorial.

INTRODUCTION: THE NEED FOR NEW MODELS

One of the main axes of research in geostatistics today is development of methods for estimating recoverable reserves. (For a comprehensive review of the state of the art, see Maréchal, 1984). Two of the better known estimation methods, disjunctive kriging in its present form and multigaussian kriging, require data to be transformed to normality. This transformation, which is usually called an anamorphosis, poses some insoluble problems when large numbers of identical values are involved, as often happens with zero grades in uranium or gold deposits. A clear description of the problem is given by Verly (1984) who presents an example with 30% zero values. In his case, a criterion had to be found for ranking 150 zeros (out of a total of 500 values). As this determines the distribution of normal scores and hence influences final estimates, this choice is critical. Verly proposes one method for ranking the zeros. But one can ask whether a better solution would not be to look for models based on probability density functions (p.d.f.) other than the normal, which could handle a spike at the origin in a more satisfactory way, rather than patch up the existing methods based on the normal distribution.

Data sets with a large spike at the origin are not the only cases where existing methods fail. They also are inappropriate for handling discrete varia-

[1]Manuscript received 10 June 1985; accepted 30 January 1986.
[2]Centre de Geostatistique et de Morphologie Mathematique, Ecole Nationale Superieure des Mines de Paris, 35 Rue Saint-Honore, 77305 Fontainebleau, France.

0882-8121/86/1100-0711$05.00/0

bles such as stone count in diamonds or grouped data that arise when size and density distributions are being measured. The standard sampling procedure for the latter is to weigh material which passes through a particular size sieve or which floats in a liquid of a specified density. Although variables are continuous in this case, the available sample data are grouped into a relatively small number of classes. They therefore present similar problems to those encountered handling discrete variables, namely that a normal transformation is inappropriate. Consequently, a better solution is to avoid methods based on the normal distribution. One might be tempted to try distribution-free methods such as indicator kriging or probability kriging. However, as Journel (1984) pointed out in his article on these methods, "Data-support geostatistics is the absolute frontier of distribution-free geostatistics"; that is, whereas these methods possibly may be useful for estimating variables with the same support as the data, they definitely ought not be used for estimating variables with a different support such as grade of selective mining units or blocks. We would question the advisability of using these methods even for variables with the same support because of the "destructuration" effect seen at richer grades. Matheron (1982) has demonstrated that shape of the variogram changes with cut-off grade. In the bivariate normal model, the variogram tends toward a pure nugget effect model. (One exception to this is the mosaic model in which the shape of the variogram is the same). In addition to these theoretical difficulties with indicator variograms at large cut-off grades, practical problems also exist. As the cut-off for an indicator increases, the indicator functions for the data contain more and more zeros and, consequently, the corresponding variograms which depend entirely on location of the few remaining values become unstable. The underlying structure of the data seems to disappear.

Consequently, we conclude that several types of variables (those with a spike at the origin, discrete variables, grouped data) are inappropriate for existing recovery estimation methods. A real need is apparent for models based on a distribution other than the normal distribution. Moreover, these must have a mathematically consistent procedure for change of support.

One possible approach is to find additional types of disjunctive kriging based on other distributions. By this, we mean that if a suitable decomposition of the bivariate distribution into orthogonal factors (which need not be polynomials) can be found, an estimate of recoverable reserves could be obtained by kriging each of these factors separately in the same way as is done in ordinary disjunctive kriging with Hermite polynomials. At present, the standard decomposition of the bivariate normal $g(x, y)$ into Hermite polynomials

$$g(x, y) = \sum C_n H_n(x) H_n(y) g(x) g(y) \tag{1}$$

is the basis for ordinary disjunctive kriging. The idea is therefore to find similar representations (which Matheron, 1976, has called isofactorial representations) for other distributions. Orthogonal factors, being uncorrelated, could be kriged independently to give estimated recovery. The literature on orthogonal polynomials (see, e.g., Szego, 1939) and on orthogonal polynomials to approximate/estimation densities (see Wertz, 1979 for a bibliography) gives many examples of families of orthogonal polynomials associated with different distributions which possibly could be used for the marginal distribution. However, the crucial problem is to find bivariate distributions which can be expressed in terms of these orthogonal factors. One always could construct likely looking "bivariate distributions" from orthogonal polynomials. But no guarantee exists that these functions would have properties required for a p.d.f. This is why models presented in this paper have been constructed from random functions. A second important reason for working with random functions is that this makes it possible to develop meaningful models for change of support.

Our objective in the long term is to present a range of models (both discrete and continuous) with these properties. Some work has been published recently. An overview of the problem and a solution for the case of the negative binomial distribution were given in Matheron (1984). A general method for constructing isofactorial distributions to suit an arbitrary marginal distribution is presented in Matheron (1985a).

Unfortunately, much early work on this subject (Matheron, 1973, 1975a, 1975b) has only been available in the Center's internal reports in French. The objective of this paper, and a following Part II, is to present a revised translation of this work. We have chosen to present an updated translation rather than just summarizing the main results because this makes it possible to follow the evolution of the development of the models and hence to see why certain techniques (infinitesimal generators and Markov semigroups) are used in the later work.

Readers may be curious why this work, which has been floating around for more than a decade, had not been published earlier. The reason is that the original paper on the subject (Matheron, 1976) gave an overview of disjunctive kriging, in general, and the normal version with Hermite polynomials, in particular. At that time, the need for other models had not yet been recognized. However, practical experience has indicated that it is needed. Several case studies using disjunctive kriging based on the gamma distribution and the negative binomial have been completed and are currently in press (Lantuéjoul, Lajaunie, 1986; Nelson and Guibal, 1986). So it is now vital to publish the isofactorial models in their general form.

The earliest work on disjunctive kriging appears in internal report N-360 (Matheron, 1973). Although the first four sections were published in English at the first N.A.T.O. Workshop, the rest has remained unpublished. In order to make these easier to understand, we review the essential results on disjunctive kriging using Hermite polynomials and the normal distribution before going on to discuss disjunctive kriging with Laguerre Polynomials and a gamma (γ) distribution.

DISJUNCTIVE KRIGING WITH HERMITE POLYNOMIAL MODELS

First, some properties of Hermite polynomials are reviewed. (Proofs can be found in any text with a chapter on orthogonal polynomials. See, for example, Szego, 1939.) These polynomials are defined by the equation

$$e^{-x^2/2} H_n(x) = \frac{d^n}{dx^n} e^{-x^2/2} \qquad n = 0, 1, 2, \ldots$$

They form a set of orthogonal polynomials with respect to the normal distribution. That is

$$\int H_n(x)\, H_m(x)\, \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx = \delta_{nm}\, n!$$

The normed polynomials $\eta_n(x) = H_n(x)/\sqrt{n!}$ form an orthonormal basis for the Hilbert space $L^2(R, e^{-x^2/2}/\sqrt{2\pi})$. The next step is to prove that the bivariate normal p.d.f, $g(x, y)$, can be expressed in terms of marginal distributions $g(.)$ and Hermite polynomials

$$g(x, y) = g(x)\, g(y) \sum \rho^n/n!(x)\, H_n(y) \tag{2}$$

This result is well-known in statistical literature (see Anderson, 1958). Matheron (1976) has proved that, provided a bivariate p.d.f. can be expressed in an isofactorial form, estimates of recovery functions [or any function in $L^2(R, g)$] can be obtained by kriging Hermite polynomials separately.

What is important in the present context is that the bivariate normal probability density function has an isofactorial representation with Hermite polynomials as the system or orthogonal polynomials and with ρ^n, $n = 0, 1, \ldots$ as the corresponding eigenvalues. (Eigenvalues appear naturally when the theory is developed in terms of projections.) We now go on to develop a similar representation for a type of bivariate γ distribution. Unlike the bivariate normal where the standard canonical form is well-known, several bivariate distributions exist with the γ as their marginal distribution. Because these are less well-known, details of the development will be given. Readers can also consult Lukacs (1977) for the multivariate γ distribution.

DISJUNCTIVE KRIGING WITH LAGUERRE POLYNOMIAL MODELS

The correlation between two regionalized variables $Z(x)$ and $Z(x + h)$, each with a γ distribution, must be determined.

Consider a random measure μ on $\mathbb{R}^n$, which is stationary and orthogonal, and has a γ distribution; that is, for all Borel sets $B \subset \mathbb{R}^n$ with volume V

$$E[e^{-\lambda\mu(B)}] = e^{-\psi(\lambda)}$$

where $\psi(\lambda) = \theta V \log(1 + \lambda)$. More information on random measures and Borel sets can be found in Yaglom (1962).

Let $Z(x)$ denote regularization of μ corresponding to a specified compact set B; i.e.

$$Z(x) = \int \mu(d\xi)\, 1_B(x + \xi)$$

where $1_B(.)$ is the indicator function for set B. The transitive covariogram of B is defined as

$$K(h) = \int 1_B(x)\, 1_B(x + h)\, dx$$

For more information on transitive covariograms, see Matheron (1965) or Serra (1982). One may show that

$$E\,[e^{-\lambda Z(x) - \nu Z(x+h)}] = \exp\{-\theta\, K(h)\, \psi(\lambda + \nu) - \theta[K(o) - K(h)][\psi(\lambda) + \psi(\nu)]\} \tag{3}$$

This is done by splitting $Z(x)$ and $Z(x + h)$ into disjoint components

$$Z(x) = \mu(B_x \cap B_{x+h}) + \mu(B_x \setminus B_x \cap B_{x+h})$$

$$Z(x + h) = \mu(B_x \cap B_{x+h}) + \mu(B_{x+h} \setminus B_x \cap B_{x+h})$$

Components are three independent γ variables with parameter values of $\theta[K(o) - K(h)]$ for two of the variables, and of $\theta\, K(h)$ for the one common to $Z(x)$ and $Z(x + h)$. Equation (3) follows from this.

So, $Z(x)$ and $Z(x + h)$ are correlated γ variables with parameter $\alpha = \theta\, K(o) = \theta\, V(B)$, where $V(B)$ is the volume of the ball. The correlation coefficient ρ between $Z(x)$ and $Z(x + h)$ is $K(h)/K(o)$. For simplicity, the scale parameter has been set to 1.

Laguerre Polynomials

The system of orthogonal polynomials relative to the γ distribution is defined by

$$(d^n/dx^n)\, x^{n+\alpha-1}\, e^{-x} = (-1)^n\, n!\, x^{\alpha-1}\, e^{-x}\, L_n(x) \qquad \alpha > 0$$

These polynomials form an orthogonal basis for the Hilbert space $L^2\,(R_+, (1/\Gamma(\alpha)x^{\alpha-1}\, e^{-x})$.

The bivariate p.d.f. of $X = Z(x)$ and $Y = Z(x + h)$ can be shown to have the following isofactorial representation

$$f(x, y) = \sum U_n \frac{L_n(x)\, L_n(y)}{\|L_n\|^2} x^{n+\alpha-1} y^{n+\alpha-1} e^{-x-y} \qquad \text{where}$$

$$U_n = \frac{\Gamma(\rho\alpha + n)\, \Gamma(\alpha)}{\Gamma(\rho\alpha)\, \Gamma(\alpha + n)} \qquad \text{(See Appendix A)} \qquad \text{and} \tag{4}$$

$$\|L_n\|^2 = \frac{\Gamma(\alpha + n)}{\Gamma(\alpha)\, n!}$$

In contrast to the normal distribution where $U_n = \rho^n$, we have

$$U_n = E\,[T^n] \tag{5}$$

where T is a β random variable $\beta\,[\rho\alpha,\ (1 - \rho)\alpha]$, with mean ρ. So although a gaussian anamorphosis would transform the marginal distribution to a normal one if it were applied to this case, the bivariate distribution would not be gaussian (normal)—but this should have been obvious from the outset.

This raises the question whether a bivariate γ distribution exists with the p.d.f.

$$f(x, y) = \rho^n \frac{L_n(x)\, L_n(y)}{\|L_n\|^2} x^{n+\alpha-1}\, y^{n+\alpha-1}\, e^{-x-y} \tag{6}$$

This is shown to be true in Appendix B. This second bivariate γ distribution represents a diffusion-type process. It is the basis for some recent work on change of support models (Matheron, 1985b).

The following section is devoted to the search for other models with polynomial factors.

LOOKING FOR OTHER MODELS WITH POLYNOMIAL FACTORS

The method used for finding bivariate distributions with an isofactorial representation suitable for disjunctive kriging is the same for both normal and γ distributions. In summary, it consists of finding random functions with the following properties

(a) $Z(x)$ must be stationary. For a given x, a complete countable set of orthonormal functions χ_n associated with the distribution $w(dz)$ of $Z(x)$ must exist. The χ_n form the basis for Hilbert space $L^2(\mathbb{R}, w)$.

(b) For any two points x and y, the joint p.d.f. of $Z(x)$ and $Z(y)$ must be of the form

$$\Phi(z, z')\, w(dz)\, w(dz') \qquad \text{where}$$

$$\Phi(z, z') = \sum U_n(x, y)\, \chi_n(z)\, \chi_n(z')$$

Several problems arise when one is looking for random functions satisfying these conditions. Because the bivariate p.d.f. $f_{\alpha\beta}(Z, Z')$ plays the same role in disjunctive kriging as covariance does in ordinary kriging, some sort of generalization of Bochner's theorem is required to characterize families of bivariate distributions which can be associated with random functions. To be more precise, for a given family $F(dz, dz'; x_1, x_2)$ of bivariate distributions with $x_1, x_2 \in \mathbb{R}^n$, what is the condition that guarantees the existence of a R.F. $Z(x)$ on $\mathbb{R}^n$ having $F(dz, dz'; x_1, x_2)$ as the bivariate distribution of $Z(x_1)$, $Z(x_2)$? Clearly for all $x \in \mathbb{R}^n$, the marginal probability $w(.) = F(., R, x, x')$ must be independent of x.

By analogy with the covariance, which must be positive definite, we could postulate a similar condition. That is, for any $x_1, \ldots, x_k \in \mathbb{R}^n$ and $f_i \in L^2(R', w_i)$ $(i = 1, \ldots, k)$, we could require that

$$E\left[\left[\sum f_i(Z(x_i))\right]^2\right] \geq 0 \qquad \text{i.e.}$$

$$\sum_i \sum_j \int f_i(Z)\, f_j(Z')\, F(dz, dz'; x_i, x_j) \geq 0 \tag{7}$$

But although this condition is necessary, it is hardly sufficient. To see this we consider a partition B_k of R. The all or nothing R.F.s $1_{B_k}[Z(x)]$ associated with a random partition A_k of $\mathbb{R}^n$

$$A_k = \{x : Z(x) \in B_k\} \qquad \text{and}$$

$$P(x \in A_k, x' \in A_{k'}) = F(B_k, B_{k'}; x, x') \tag{8}$$

Condition (7) means that covariances $C_{kk'}(x, x') = F(B_k, B_{k'}; x, x')$ form the covariance matrix of a vectorial R.F. $(Y_1(x), \ldots, Y_k(x)]$ but no reason exists to think that this vectorial R.F. need necessarily be the indicator function of a random partition.

However, this suggests that the following condition should be necessary and sufficient: that for every countable partition B_k, a random partition A_k satisfying Eq. 8 exists.

This draws our attention to the following problem. Under what conditions is a family $C_{kk'}(x, x')$ the covariance matrix of a random partition? One particular, simpler case of this is: under what conditions is the function $C(x, x')$ the covariance of a random set?

This problem is by no means trivial. For example, covariances of the form $\exp\{-|x - x'|^2\}$ cannot be associated with any random set. (As they have a second derivative, they have to represent a mean square differentiable R.F.).

Because these problems are difficult, we go on to attack the question from a different point of view. When dealing with ordinary variograms and covariances, the difficulty of testing whether a given function is positive definite and hence could be used as a covariance, meant that new covariance models usually are developed by construction from a known regionalized variable. Similarly the easiest way to produce bivariate distributions with the required properties is by regularizing a stationary orthogonal measure μ e.g.

$$Z(x) = \int k(x + \xi)\, \mu(d\xi)$$

where k is given.

Suppose that the (infinitely divisible) distribution associated with μ is defined by

$$E\,[e^{-\lambda\mu(V)}] = e^{V\psi(\lambda)}$$

then, using the same reasoning as for the γ distribution, $E\,[e^{-\lambda Z(x) - \nu Z(x+h)}]$ can be shown to be equal to $\exp\,\{[K(o) - K(h)]\,[\psi(\lambda) + \psi(\nu)] + K(h)\,\psi(\lambda + \nu)\}$ where $K(h)$ is the transitive covariogram. Substituting $\Phi(\lambda) = \exp\,\{K(o)\,\psi(\lambda)\}$ and $\rho = K(h)/K(o)$ gives

$$\Phi(\lambda, \nu) = \Phi(\lambda)^{1-\rho}\,\Phi(\nu)^{1-\rho}\,\Phi(\lambda + \nu)^{\rho} \tag{9}$$

Models with Polynomial Factors

The next step is to find distributions satisfying eq. 9 and having polynomial factors χ_n. Insistence on having polynomial factors is because they are easy to compute. A more general method (Matheron, 1985a) now has been developed. In the preceding sections, the normal distribution and the γ distribution were shown to have these properties. We now go on to show that the same is true of the Poisson distribution and the negative binomial. Moreover, these are the only nontrivial distributions which satisfy eq. 9.

If a distribution $w(dx)$ has a set of polynomials χ_n which form a basis for Hilbert space $L^2(\mathbb{R}^n, w)$, the following properties can be shown to be equivalent to ensure that the bivariate distribution of X and Y has a symmetric p.d.f. of the form

$$\Phi(x, y)\; w(dx)\; w(dy)$$

1. $\Phi(x, y)$ is of the form $\Sigma\, U_n\, \chi_n(x)\, \chi_n(y)$
2. For all $n \geq 0$, $E[X^n | Y]$ is a polynomial of degree n in y

Proof is given in Appendix C. This can be used to show that the Laplace transform $\theta(\lambda, \nu)$ of a bivariate distribution with an isofactorial representation satisfies

$$\theta(\lambda)^{1-\rho}\,\frac{\partial^n}{\partial\lambda^n}\,[\theta(\lambda)]^{\rho} = \sum_0^n A_{n,k}\,\frac{\partial^k}{\partial\lambda^k}\,\theta(\lambda) \qquad \text{for all } n$$

Proof is given in Appendix D.

For $n = 1$, this is of the form

$$\rho\theta' = a_1\theta + b_1\theta'$$

This always can be satisfied by putting $a_1 = 0$ and $b_1 = \rho$. For $n = 2$, put $\theta = 2e^{\psi(\lambda)}$ to obtain a differential equation of the form

$$\psi'' = a\psi'^2 + b\psi' + c$$

If $a = b = 0$, ψ is a second-degree polynomial, which leads to the normal distribution. If $a = 0$ but $b \neq 0$, $\psi(\lambda) = Ae^{b\lambda} + B$; this corresponds to the Poisson distribution.

If $a \neq 0$, the quadratic $ax^2 + bx + c$ can have 0, 1, or 2 real roots. If one real root α exists

$$\psi'' = a(\psi' - \alpha)^2$$

The solution is $\psi = \alpha\lambda - (1/a) \log (1 + c\lambda)$, which corresponds to a γ distribution (possibly translated and transposed).

If the quadratic has no real roots

$$\psi'' = a\,[(\psi' + \alpha)^2 + b^2] \qquad \text{and}$$

$$\psi(\lambda) = (1/a) \log \cos (ab\lambda + c) - \alpha\lambda - c'$$

No distribution has this as its Laplace transform. Last is the case where two (different) real roots exist

$$\psi'' = a(\psi' - \alpha)(\psi' - \beta)$$

This gives

$$\frac{\psi' - \alpha}{\psi' - \beta} = b\, e^{a(\beta - \alpha)\lambda} \qquad \text{or}$$

$$\psi'(\lambda) = \frac{\alpha - \beta b\, e^{a(\beta - \alpha)\lambda}}{1 - b\, e^{a(\beta - \alpha)\lambda}}$$

Supposing that $a(\beta - \alpha) > 0$, which is permissible because α and β always can be reversed, the solution is

$$\psi(\lambda) = \beta\lambda - \frac{1}{a} \log \left(\frac{b - e^{-a(\beta - \alpha)\lambda}}{b - 1}\right)$$

This corresponds to a negative binomial (at least up to a linear transformation), if $|b| > 1$.

So we see that in addition to the γ and normal distributions, two others (Poisson and negative binomial) also have properties needed for disjunctive kriging, at least for $n = 2$.

Poisson Distribution

The next step is to check that these properties hold for $n > 2$ for the Poisson distribution. Because this is a discrete distribution, the generating function $G(s)$ will be used instead of Laplace transforms. For the Poisson distribution with parameter θ

$$G(s) = e^{\theta(s-1)}$$

The bivariate distribution associated with eq. 9 is defined by

$$\begin{aligned} G(s, t) &= \sum P_{nn} s^n t^m \\ &= G(s)^{1-\rho}\, G(t)^{1-\rho}\, G(st)^{\rho} \end{aligned}$$

Hence

$$\log G(s, t) = (1 - \rho)\,[\log G(s) + \log G(t)] + \rho \log G(st)$$

Substituting for $G(s)$ etc. gives

$$\begin{aligned} \log G(s, t) &= (1 - \rho)\,\theta(s - 1 + t - 1) + \rho\theta(st - 1) \\ &= \theta(t - 1) + \theta(1 + \rho t - \rho)(s - 1) \end{aligned}$$

Consequently for a fixed value of n

$$\sum_n P_{nm} t^m = e^{-\theta} e^{\theta(t-1)(1-\rho)} \frac{\theta^n}{n!} (1 + \rho t - \rho)^n$$

So, the generating function $G_n(t)$ of Y for a fixed value of $X = n$ is

$$G_n(t) = (1 + \rho t - \rho)^n\, e^{\theta(1-\rho)(t-1)}$$

Conditional moments can be found by differentiating $k + 1$ times and putting $t = 0$.

$$E\,[Y(Y - 1), \ldots, (Y - k)|n] = \theta^{k+1} + \cdots + \rho^k\, n(n - 1)(n - k)$$

Thus, $E[Y^n|X]$ is indeed a polynomial of degree n in x. The eigenvalue U_n associated with polynomial χ_n is $U_n = \rho^n$, as was the case for the normal.

Last, the expression for the orthogonal polynomials associated with the Poisson distribution is

$$P_n(x) = 1 - \binom{n}{1}\frac{x}{\theta} + \binom{n}{2}\frac{x(x - 1)}{\theta^2} - \cdots + (-1)^n \frac{x, \ldots, (x - n + 1)}{\theta^n}$$

and normed polynomials are

$$w_n(x) = \sqrt{\frac{\theta^n}{n!}}\, P_n(x)$$

Negative Binomial Distribution

As with the Poisson distribution, we must show that the conditions are satisfied for $n > 2$. The negative binomial has the following generating function

$$G(s) = \left(\frac{1-\alpha}{1-\alpha s}\right)^{\beta}$$
$$= (1-\alpha)^{\beta} \frac{\sum \alpha^n \Gamma(\beta + n) s^n}{n!\Gamma(\beta)} \qquad \begin{matrix} 0 < \alpha < 1 \\ 0 < \beta \end{matrix}$$

The corresponding bivariate distribution is defined by

$$G(s, t) = E(s^X t^Y)$$
$$= G(s)^{1-\rho}\, G(t)^{1-\rho}\, G(st)^{\rho}$$

To obtain the conditional distribution of Y for a fixed value of X, random variables are split into three independent components

$$X = X_1 + Z$$
$$Y = Y_1 + Z$$

such that

$$G(s)^{1-\rho} = E(s^{X_1}) = E(s^{Y_1})$$
$$G(s)^{\rho} = E(s^Z)$$

X_1, Y_1, and Z have negative binomial distributions and

$$P(X_1 = k, Z = p) = \frac{(1-\alpha)^{\beta}\, \alpha^{k+p}\, \Gamma(\beta\rho + p)\, \Gamma[(1-\rho)\,\beta + k]}{p!\; \Gamma(\rho\beta) k!\; \Gamma[(1-\rho)\beta]}$$

The distribution of Z for a fixed value of $X = X_1 + Z$ can be deduced from

$$P(Z = p | X = n) = \frac{P(Z = p, X_1 = n - p)}{P(X = n)}$$
$$= \binom{n}{p} \frac{\Gamma(\beta)}{\Gamma(\beta\rho)\, \Gamma[(1-\rho)\beta]} \frac{\Gamma(\rho\beta + p)\, \Gamma[(1-\rho)\,\beta + n - p]}{\Gamma(\beta + n)}$$
$$\text{for } 0 \le p \le n$$

This can be shown to be a binomial distribution with parameters n and p where the value of p is chosen at random from a β distribution.

(To see this, let p be a binomial variable with parameters n and x and let X be a r.v. with a β distribution with parameters $\rho\beta$ and $(1 - \rho)\beta$. The p.d.f. of the β distribution is

$$f(x) = \frac{\Gamma(\beta)}{\Gamma(\rho\beta)\,\Gamma[(1-\rho)\beta]}\, x^{\rho\beta-1}\,(1-x)^{(1-\rho)\beta-1}$$

$$P(K = k|x) = \binom{n}{p} x^k\,(1-x)^{n-k}$$

$$P(K = k) = \int_0^1 P(K = k|x)\, f(x)\, dx$$

Substituting expressions for $P(K = k|x)$ and $f(x)$ into this equation and integrating gives the required relation).

The generating function of Z given X can be deduced from this

$$E(s^Z|X = n) = \sum_p C\binom{n}{p} \int s^p x^p\,(1-x)^{n-p}\, x^{\rho\beta-1}\,(1-x)^{(1-\rho)\beta-1}\, dx \qquad \text{where}$$

$$C = \frac{\Gamma(\beta)}{\Gamma(\rho\beta)\,\Gamma[(1-\rho)\beta]}$$

Reversing the order of integration and summation and noting that

$$\sum \binom{n}{p} (sx)^p\,(1-x)^{n-p} = (1 - x + sx)^n$$

leads to

$$E(s^Z|X = n) = C \int_0^1 (1 - x + sx)^n\, x^{\rho\beta}\,(1-x)^{(1-\rho)\beta-1}\, dx$$

Consequently the generating function of $Y = Z + Y_1$ is

$$E\,[s^Y|X = n] = \left(\frac{(1-\alpha)}{(1-\alpha s)}\right)^{(1-\rho)\beta} E\,[s^Z|n]$$

Differentiating the conditional generating function for Z, p times and putting $s = 1$, gives

$$E\,[Z(Z-1)\cdots(Z-p+1)|X = n]$$
$$= n(n-1)\cdots(n-p+1)\int_0 x^p f_\beta(x)\, dx$$

where $f_\beta(x)$ is the p.d.f. of a β distribution. This shows that $E[Z^p|X = n]$ is a polynomial of degree p in n and hence so is $E[Y^p|X = n]$.

Therefore, the negative binomial satisfies the requirements stated earlier. The explicit expression for the orthogonal polynomials will not be given here. However, we note that their eigenvalues are given by

$$U_p = \int_0^1 x^p f_\beta(x)\, dx = \frac{\Gamma(\rho\beta + p)\,\Gamma(\beta)}{\Gamma(\rho\beta)\,\Gamma(\beta + p)}$$

CONCLUSION

The objective of this paper is to show that disjunctive kriging can be used with distributions other than the Gaussian (normal) distribution. In this paper, three particular examples (Poisson distribution, negative binomial, and γ) have been presented. Although these methods were developed as early as 1973, the need for them has only been recognized recently. Several case studies using these methods have been completed now and others are under way. Results of some studies are confidential (which is often the case when dealing with precious substances). However, a series of comparative case studies on a uranium deposit are currently in press. Sans (1986) carried out a detailed comparison of the actual production figures with the estimated grade/tonnage curves obtained using the discretized Gaussian model. As is often the case with uranium deposits, a significant spike of zero values at the origin occurs. This suggested that it may be worthwhile trying other types of models. Lantuéjoul and Lajaunie (1986) used new isofactorial models based on the γ distribution and the negative binomial to estimate global recoverable reserves. Their paper describes problems encountered when putting these methods into practice—for example, how to do the anamorphosis.

The fact that several case studies have been carried out using isofactorial models other than the usual Hermite polynomial/normal distribution highlights the need for a paper detailing the underlying theory.

In two other papers from the mid-1970s, Matheron (1975a and b) showed that other distributions can be described with polynomial factors, but unlike the four considered here, the others are not infinitely divisible and hence do not satisfy (9). This work, which will be presented in a separate paper, was developed using a different approach—that of infinitesimal generators.

APPENDIX A

We show that the bivariate γ distribution defined by regularizing a random measure with a γ distribution has an isofactorial representation.

The regionalized variable $Z(x)$ was defined as

$$Z(x) = \int \mu d(\xi)\; 1_B(x + \xi)$$

where μ is a stationary orthogonal random measure with a γ distribution, and B is a ball. The expansion of $e^{-\lambda x}$ in terms of Laguerre polynomials is

$$e^{-\lambda x} = \sum (-1)^n \frac{\lambda^n}{(1 + \lambda)^{n+\alpha}} L_n(x)$$

Using this, we see that if the bivariate distribution has an isofactorial representation

$$E[e^{-\lambda X - \nu Y}] = \sum U_n \frac{\lambda^n}{(1 + \lambda)^{\alpha+n}} \frac{\nu^n}{(1 + \nu)^{\alpha+n}} \cdot \frac{\Gamma(\alpha + n)}{n!\,\Gamma(\alpha)} \qquad \text{(A1)}$$

A second expression for $E[e^{-\lambda x - \nu Y}]$ can be obtained by calculating the Laplace transform directly from the decomposition of $Z(x) = X$ and $Z(x + h) = Y$ into three disjoint components

$$E[e^{-\lambda X - \nu Y}] = \frac{1}{(1 + \lambda)^{\alpha} (1 + \nu)^{\alpha}} \left[1 - \frac{\lambda\nu}{(1 + \lambda)(1 + \nu)}\right]^{-\rho\alpha}$$

The last term can be expanded as a negative binomial; viz

$$\left[1 - \frac{\lambda\nu}{(1 + \lambda)(1 + \nu)}\right]^{-\rho\alpha} = \sum \frac{\Gamma(\rho\alpha + n)}{n!\,\Gamma(\rho\alpha)} \cdot \left[\frac{\lambda\nu}{(1 + \lambda)(1 + \nu)}\right]^n$$

Hence

$$E[e^{-\lambda X} e^{-\nu Y}] = \sum \frac{\Gamma(\rho\alpha + n)}{n!\,\Gamma(\rho\alpha)} \frac{\lambda^n \nu^n}{(1 + \lambda)^{n+\alpha} (1 + \nu)^{n+\alpha}}$$

Comparing these two expressions for the Laplace transform shows that factorial representation (A1) is valid and that eigenvalues are

$$U_n = \frac{\Gamma(\rho\alpha + n)}{\Gamma(\rho\alpha)} \frac{\Gamma(\alpha)}{\Gamma(\alpha + n)}$$

So we obtain

$$f(x, y) = \sum U_n L_n(x) L_n(y) x^{n+\alpha-1} y^{n+\alpha-1} e^{-x-y}$$

APPENDIX B

We wish to show that, in fact, a bivariate γ distribution exists whose p.d.f. is

$$f(x, y) = \sum \rho^n L_n(x) L_n(y) x^{n+\alpha-1} y^{n+\alpha-1} e^{-x-y} \qquad \text{(B1)}$$

From (A1), the Laplace transform of $f(x, y)$ is

$$\Phi(\lambda, \nu) = \sum \frac{\rho^n}{n!} \frac{\Gamma(n + \alpha)}{\Gamma(\alpha)} \frac{\lambda^n \nu^n}{(1 + \lambda)^{n+\alpha} (1 + \nu)^{n+\alpha}}$$

Remembering the form for the negative binomial expansion, we see that

$$\Phi(\lambda, \nu) = \frac{1}{[(1 + \lambda)(1 + \nu)]^{\alpha}} \left[1 - \frac{\rho\lambda\nu}{(1 + \lambda)(1 + \nu)}\right]^{-\alpha}$$
$$= [1 + \lambda + \nu + (1 - \rho)\lambda\nu]^{-\alpha}$$

We now have to check that this Laplace transform does correspond to some distribution (an infinitely divisible one, at that). To do this we consider the following expansion

$$\Phi(\lambda, \nu) = \frac{(1 - \rho)^{\alpha}}{[1 + \lambda(1 - \rho)]^{\alpha}\,[1 + \mu(1 - \rho)]^{a}}$$
$$\cdot \left\{1 - \frac{\rho}{[1 + \lambda(1 - \rho)]\,[1 + \nu(1 - \rho)]}\right\}^{-\alpha}$$
$$= \frac{(1 - \rho)^{\alpha}}{\Gamma(\alpha)} \sum \frac{\Gamma(\alpha + n)}{n!} \frac{\rho^{n}}{[1 + \lambda(1 - \rho)]^{\alpha + n}\,[1 + \nu(1 - \rho)]^{\alpha + n}}$$

This is the Laplace transform of a mixture of two γ distributions with parameters $\alpha + n$ where the parameter n of the γ distributions varies according to a negative binomial distribution with p.d.f.

$$P(N = n) = \frac{\Gamma(n + \alpha)}{n!\,\Gamma(\alpha)} \rho^{n} (1 - \rho)^{\alpha}$$

After an easy calculation we obtain

$$f(x, y) = \sum_{0}^{\infty} \frac{\Gamma(n + \alpha)}{n!\,\Gamma(\alpha)} \rho^{n} \frac{(xy)^{\alpha + n - 1}}{(1 - \rho)^{2n + \alpha}} e^{-(x + y)/(1 - \rho)} \tag{6}$$

The Laplace transform given earlier can be seen to correspond to this p.d.f.

APPENDIX C

Theorem: If a distribution $w(dx)$ has a set of polynomials χ_n which form a basis for Hilbert space $L^2(\mathbb{R}^n, W)$, the following properties are equivalent to insure that the bivariate distribution of X and Y has a symmetric p.d.f. of the form

$$\Phi(x, y)\, w(dx)\, w(dy)$$

1. $\Phi(x, y)$ is of the form $\Sigma\, U_n\, \chi_n(x)\, \chi_n(y)$
2. For all $n \geqq 0$, $E\,[X^n|Y]$ is a polynomial of degree n in y

Proof. If (1) holds, then

$$E\,[X^n|Y] = \sum_{p} U_p\, \chi_p(y) \int \chi_p(x)\, x^n\, w(dx)$$

Because $\int \chi_p(x)\, x^n\, w(dx) = 0$ for all $p > n$

$$E\,[X^n | Y = y] = \sum_{p \leq n} U_p\, \langle x^n, \chi_p \rangle\, \chi_p(y)$$

where $\langle x^n, \chi_p \rangle$ denotes $\int \chi_p(x)\, x^n\, w(dx)$.

This expression is a polynomial of degree $\leq n$. In fact, it is precisely of degree n, because $\langle x^n, \chi_p \rangle$ cannot be zero because

$$x^n = \sum_{p=0}^{n} \langle x^n, \chi_p \rangle\, \chi_p(x)$$

Conversely, suppose (2) holds. We have to show that orthogonal polynomials $\chi_n(x)$ satisfy

$$E\,[\chi_n(x) | y] = U_n\, \chi_n(y)$$

The space $\mathcal{P}_{n+1}$ of dimension $n + 1$ made up of polynomials of degree $\leq n$ is invariant under the operator $E[X|Y]$. Moreover this operator maps $\mathcal{P}_{n+1}$ onto itself, because the degree of the polynomials is conserved. Consequently, this symmetric operator has $n + 1$ orthogonal eigenvectors in $\mathcal{P}_{n+1}$, which are just χ_k, $k = 0, 1, \ldots, n$. So (1) follows.

The next step is to look for conditions under which distributions of form (9) have polynomial factors. From the discussion given above, it is clear that (2) must hold.

APPENDIX D

To show that Laplace transform $\theta(\lambda, \nu)$ of a bivariate distribution with a symmetric p.d.f. of the form $\Phi(x, y)\, w(dx)\, w(dy)$ satisfies

$$\theta(\lambda)^{1-\rho} \frac{\partial^n}{\partial \lambda^n} [\theta(\lambda)]^\rho = \sum_0^n A_{n,k} \frac{\partial^k}{\partial \lambda^k} \theta(\lambda)$$

First we remind readers that the Laplace transform of a distribution is another name for the moment generating function and that the expectation $E(X^n)$ can be obtained from the Laplace transform for the distribution of X as follows

$$E(X^n) = \frac{\partial^n}{\partial \lambda^n} \theta_X(\lambda)|_{\lambda=0}\, (-1)^n$$

Similarly, the conditional expectation $E(X^n | Y = y)$ satisfies the relation

$$\int e^{-\lambda y}\, E[X^n | Y = y]\, w(dy) = (-1)^n\, (\partial^n / \partial \nu^n)\, \theta(\lambda, \nu)|_{\nu=0}$$

If we now apply the preceding theorem, we see that the Laplace transform θ must satisfy

$$\theta(\lambda)^{1-\rho} \frac{\partial^n}{\partial\lambda^n} [\theta(\lambda)]^\rho = \sum_0^n A_{n,k} \frac{\partial^k}{\partial\lambda^k} \theta(\lambda)$$

for all n.

REFERENCES

Anderson, T. W., 1958, Introduction to multivariate statistical analysis: Wiley & Sons, New York.

Journel, A., 1984, The place of non-parametric geostatistics: *in* G. Verly (Ed.), Geostatistics for natural resource characterization, Proceedings NATO A.S.I.: Reidel Publishing Co., Dordrecht, p. 307–336.

Lantuéjoul, C., Lajaunie, C., and Rivoirard, J., 1986, Testing several change of support models used to estimate global reserves; *in* G. Matheron and M. Armstrong (Eds.), Geostatistical case-studies: Reidel Publishing Co, Dordrecht, in press.

Lukacs, F., 1977, Multivariate Analysis IV p. 119 publisher, etc.

Maréchal, A., 1984, Recovery estimation: a review of models and methods: *in* G. Verly (Ed.), Geostatistics for natural resources characterization, Proceedings NATO A.S.I.: Reidel Publishing Co, Dordrecht, p. 385–420.

Matheron, G., 1970, La theorie des variables regionalisees et ses applications: Ed Ecole National Superieure des Mines de Paris, 212 p.

Matheron, G., 1973, Le krigeage disjonctif: Internal note N-360; Centre de Geostatistique, Fontainebleau, 40 p.

Matheron, G, 1975a, Complements sur les modeles isofactoriels: Internal note N-432; Centre de Geostatistique, Fontainebleau, 21 p.

Matheron, G., 1975b, Modeles isofactoriels discrets et modeles de Walsh: Internal note N-449; Centre de Geostatistique, Fontainebleau, 24 p.

Matheron, G., 1976, A simple substitute for conditional expectation: The disjunctive kriging; *in* M. Guarascio et al. (Eds.), Advanced geostatistics in the mining industry, Proceedings NATO A.S.I.: Reidel Publishing Co, Dordrecht, p. 221–236.

Matheron, G., 1982, La destructuration des hautes teneurs et le krigeage des indicatrices; Internal note N-761; Centre de Geostatistique, Fontainebleau, 33 p.

Matheron, G., 1984, Isofactorial Models and Change of Support, *in* G. Verly (Ed.), Geostatistics for natural resources characterization, Proceedings NATO A.S.I.: Reidel Publishing Co, Dordrecht, p. 449–468.

Matheron, G., 1985a, Une methodologie generale pour les modeles isofactoriels discrets: Sci. Terre. Inform., v. 21, p. 1–78.

Matheron, G., 1985b, Change of support for diffusion-type random functions: Math. Geol., v. 15, no. 2, p. 137–166.

Nelson, G. and Guibal, D., 1986, Recoverable reserves on an Australian gold deposit, *in* G. Matheron and M. Armstrong (Eds.), Geostatistical case-studies: Reidel Publishing Co, Dordrecht, in press.

Sans, H., 1986, Comparing estimates to real uranium grades: The geostatistical approach; *in* G. Matheron and M. Armstrong (Eds.), Geostatistical case-studies, Reidel Publishing Co, Dordrecht, in press.

Serra, J., 1982, Image analysis and mathematical morphology: Academic Press, London, 610 p.
Szego, G., 1939, Orthogonal polynomials: American Mathematical Society Colloquium Publications, v. XXIII, Providence, Rhode Island, U.S.A., 420 p.
Verly, G., 1984, The block distribution given a point multivariate normal distribution; *in* G. Verly (Ed.), Geostatistics for natural resources characterization, Proceedings NATO A.S.I.: Reidel Publishing Co, Dordrecht. p. 495–516.
Wertz, W., 1979, Statistical density estimation—A bibliography: Int. Stat. Rev., p. 155–175.
Yaglom, A. M., 1962, Stationary Random Functions: Dover Publications, New York, 253 p.