FONTAINEBLEAU/CG C-131

ADVANCED GEOSTATISTICS

(Cours Afrique du Sud)

J. RIVOIRARD *Octobre 1987*

ADVANCED GEOSTATISTICS

CONTENTS

GEOSTATISTICS FOR SKEW DISTRIBUTIONS

INTRODUCTION

The presence of a small percentage of high grades usually linked with a short-range structure makes it difficult to estimate skewly distributed grades, since all differences or errors that occur are magnified. Sometimes new geostatistical models are needed but the general concepts remain the same. Above all, difficulties do arise when using classical tools. This requires great care.

This course is designed for practioners who are already familiar with basic geostatistics. These lecture notes merely provide an overview of each topic.

SERVICE VARIABLES

1.WIDTH AND ACCUMULATION (2D)

The mean grade of a mineralized layer, defined as the ratio metal/ore is the average of the grades weighted by the width. This is why we need to use the two service variables: width and accumulation. These can be averaged directly to give the mean width and the mean accumulation. The ratio of these two is the mean grade.

2.ORE AS AN INDICATOR FUNCTION AND METAL

When the mineralization is present within formations whose limits ($Z(x) = z_0$) are unknown, it is convenient to represent the presence or absence of the mineralization at a point x using the indicator function:

$$1_{Z(x)\geq z_0} \quad \begin{array}{ll} = 1 & \text{if } Z(x) \geq z_0 \\ = 0 & \text{otherwise} \end{array} \qquad (1)$$

The mean grade of the mineralization in a block V is the mean of the grades Z(x) weighted by this indicator. So it is the ratio of the mean value of the metal:

$$Z(x)\ 1_{Z(x)\geq z_0} = \text{Metal} \qquad (2)$$

and that of the ore:

$$1_{Z(x)\geq z_0} = \text{Ore} \qquad (3)$$

Please note that the proportion of ore within the block V is just:

$$\frac{1}{V}\int_V 1_{Z(x)\geq z_0}\,dx \qquad (4)$$

3.ESTIMATION OF THE SERVICE VARIABLES

The service variable representing the ore is normally much more structured than the other service variable (the metal). When these two variables are kriged separately, each of the samples has different weights for each estimator. For instance a rich sample may have a large weight for one variable and a small one for the other, causing the ratio - the estimated grade - to be unrealistic. A more consistent relationship between the estimates of the ore, metal and grade can be obtained by cokriging the service variables.

Quite often the cross-variogram between the ore tonnage $T(x)$ and the metal quantity $Q(x)$ has the same shape as the ore variogram (which is the more regular of the two) up to a multiplicative factor p. This means that the metal $Q(x)$ can be written as:

$$Q(x) = pT(x) + R(x) + \text{constant} \tag{5}$$

where $R(x)$ is uncorrelated with the metal $T(x+h)$ for all h. In the stationary case p is the slope of the linear regression of $Q(x)$ on $T(x)$, and $R(x)$ is the residual. Kriging $T(x)$ and $R(x)$ separately is then equivalent to cokriging them, and so we get the cokriged value of $Q(x)$ from Eq(5).

SELECTION ON GRADE ESTIMATES

Selection on grade estimates is likely to reflect the qualities of grade sampling and estimation.

Let Z_v denote the real grade of the selection block v, and let Z_v^* be the estimator of Z_v. For example, it could be

- the mean of the samples inside v:

  $$\Sigma\ Z(x_i)/n$$

- a weighted average of both inside and outside samples (e.g. the kriged estimate):

  $$\Sigma\ \lambda_i\ Z(x_i) \quad \text{with } \Sigma\lambda_i = 100\%$$

1.THE BIAS $E(Z_v - Z_v^*)$

The bias is the mean difference between the true and estimated values taken <u>over all blocks without any selection</u>. Samples may not represent the extracted volume because

- part of this volume is waste and has not been sampled (e.g. the fill, or the footwall and the hanging wall of a vein),
- of sample bias (poor recovery of cuttings with, for example, loss of fine particles which are often rich in metal),
- of bias in chemical analyses.

It should be noted that a significant difference on average between Z_v and Z_v^* <u>after selection</u> does not imply the existence of such a bias.

From now on, we suppose that there is no overall bias, that is:

$$E(Z_v - Z_v^*) = 0 \qquad (6)$$

2.CONDITIONAL BIAS $E(Z_v - Z_v^* | Z_v^*)$

This occurs when the blocks estimated at $Z_v^* = z$, in fact have a different average grade: $E(Z_v^* | Z_v^* = z)$. In this case the blocks above the cutoff grade z_c have an average grade of $E(Z_v | Z_v^* \geqslant z_c)$, which is different from their estimated mean. Then the conditional bias

$$E(Z_v - Z_v^* | Z_v^*)$$

is responsible for a bias on the selected grade: $E(Z_v - Z_v^* | Z_v^* \geqslant z_c)$.

This is the mean of the conditional bias for the selected blocks. Suppose that the regression $E(Z_v|Z_v^*)$ is linear:

$$\frac{E(Z_v|Z_v^*) - m}{\sigma_v} = \frac{\rho\ (Z_v^* - m)}{\sigma_v^*} \qquad (7)$$

where $m = E(Z_v) = E(Z_v^*)$ (8)

$\sigma^2_v = Var(Z_v)$

$\sigma^2_v{}^* = Var(Z_v^*)$

ρ = correlation (Z_v, Z_v^*)

Consequently,

$$E(Z_v|Z_v^*) - m = p\ (Z_v^* - m) \qquad (9)$$

where the slope of the regression p is given by

$$p = \rho\ \sigma_v/\sigma_v^* = Cov(Z_v, Z_v^*)/Var(Z_v^*) \qquad (9b)$$

<u>Remark</u>: When there is no conditional bias, the regression is of course linear with a slope of 1. So a value of p different to 1.0 implies a conditional bias. Hence the importance, when kriging, of this parameter which can be calculated from the covariance function.

From Eq (9), we can see that

$$\underbrace{E(Z_v|Z_v^* \geqslant z_c) - m}_{\text{actual grade increase}} = p\ \underbrace{[E(Z_v^*|Z_v^* \geqslant z_c) - m]}_{\text{grade increase expected from the selection}} \qquad (10)$$

<u>If $p \leq 1$</u>, the recovered grade is always less than the predicted. The estimator Z_v^* overestimates the selected grade. This is commonly the situation when Z_v^* is the average of samples inside v since it is more variable than Z_v:

$$\sigma^2_v{}^* > \sigma^2_v \quad ==> \quad p = \rho\ \sigma_v/\sigma_v^* < 1.0 \qquad (11)$$

Note that a partial but unbiased recovery of samples makes $\sigma^2_v{}^*$ increase and hence p decrease. This highlights the importance of the sampling quality.

<u>If $p \geq 1$</u>, the recovered grade is always above the estimated one, and thus underestimated. This underestimation can lead to wrongly deciding that the deposit is uneconomic.

Particular Cases:

(i) Simple Kriging (i.e. kriging with a known mean)

From the equations p is equal to 1. So, if the regression is linear, there is no conditional bias.

(ii) Ordinary Kriging (i.e. Kriging with an unknown mean)

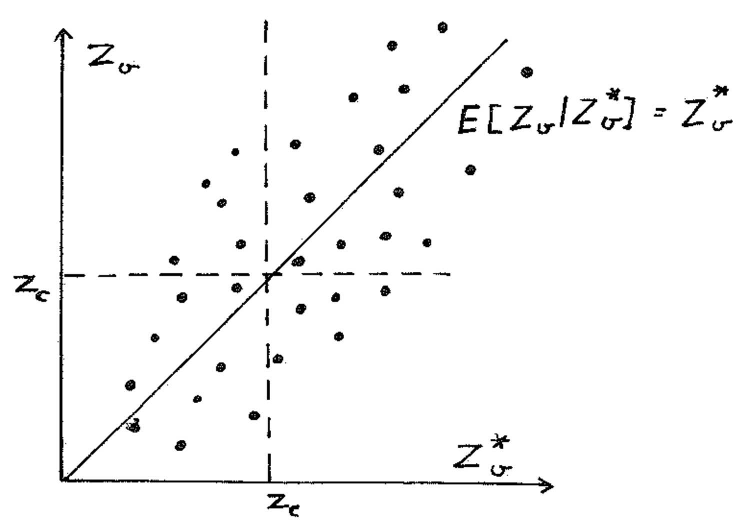
Here the known mean replaced by its locally estimated value. If this is imprecise (because the kriging neighbourhood is too small) and if the weight of the mean is not negligible, then $p<1$ and the recoverable reserves are overestimated.

Important Remark

A cutoff z_c applied on the estimated grades $Z_v{}^*$ guarantees that the expected grade $E(Z_v|Z_v{}^*)$ will be above $E(Z_v|Z_v{}^*=z_c) = h(z_c)$, but not that it will exceed z_c. This shows how illusory a cutoff is when applied to a conditionally biased estimate. To obtain a correct selection at the cutoff z_c the cutoff should be made on $E(Z_v|Z_v{}^*)$ which is $Z_v{}^*$ corrected for the conditional bias.

We now assume that there is no conditional bias.

3.ORE-WASTE ERRORS



When Z_v is unknown, selecting on the estimated grade $Z_v{}^*$ leads to:

- taking poor blocks that have been estimated rich

- leaving rich blocks that have been estimated poor

These classification errors correspond to the vertical dispersion of the scatter diagram around the bisectrice: $Var(Z_v - Z_v{}^*|Z_v{}^*)$.

Remark: Let T denote the selected tonnage, i.e. the proportion of selected blocks. These contain less metal than the really richest T% of the blocks since some of the really rich blocks have been effectively replaced by poorer ones. This is called the information effect.

4.THE ESTIMATION VARIANCE: $Var(Z_v - Z_v^*)$

We can write $Z_v - Z_v^*$ as

$$Z_v - Z_v^* = [Z_v - E(Z_v|Z_v^*)] + [E(Z_v|Z_v^*) - Z_v^*] \qquad (12)$$

where the second term $[E(Z_v|Z_v^*) - Z_v^*]$, as a function of Z_v^*, is uncorrelated with the first term.

Consequently,

$$Var(Z_v - Z_v^*)$$
$$= Var[Z_v - E(Z_v|Z_v^*)] + Var[E(Z_v|Z_v^*) - Z_v^*] \qquad (13)$$

The estimation variance is the sum of two terms. The second term corresponds to the conditional bias; while the first corresponds to the vertical dispersion of the scatter diagram around the regression (ore/waste errors, after correction of the conditional bias).

So minimizing the estimation variance is a compromise between minimizing the difference between the predicted and the recovered, and minimizing the ore/waste errors.

GRADE/TONNAGE CURVES

SUPPORT AND INFORMATION EFFECTS

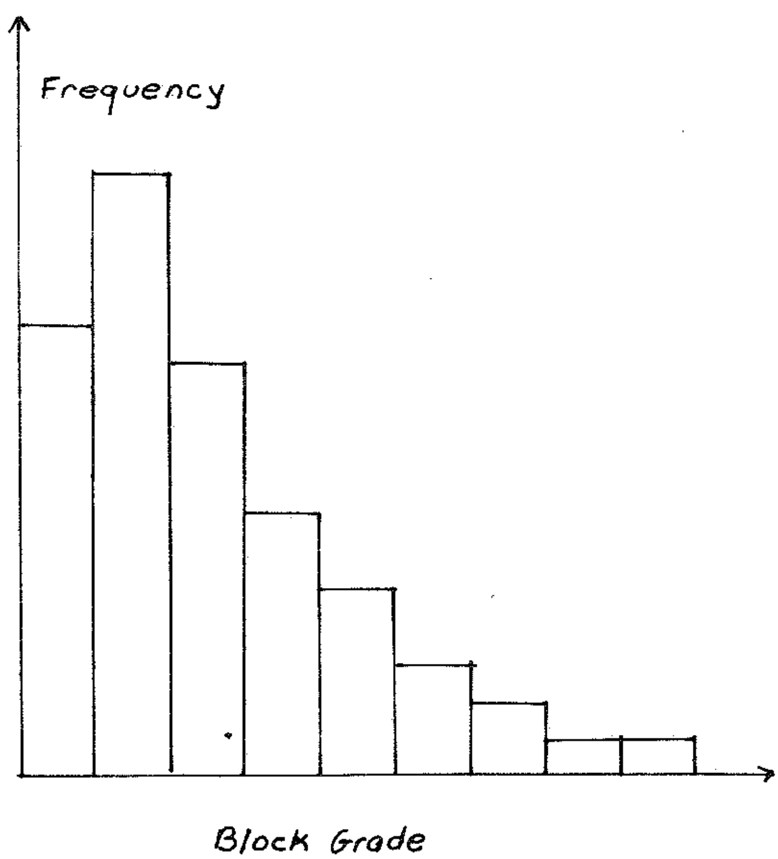
1.GRADE TONNAGE CURVES

Let us consider a deposit, or part of one, which has a constant density and contains a total tonnage T_0 at a mean grade of m. Clearly it contains T_0m tons of metal.

We now suppose that it is divided into n small blocks v_i with a constant size v. Each of these blocks contains a tonnage T_0/n. Let its grade be $Z(v_i)$. Then its metal content is $T_0Z(v_i)/n$.

Clearly

$$m = \Sigma Z(v_i)/n \qquad (13)$$



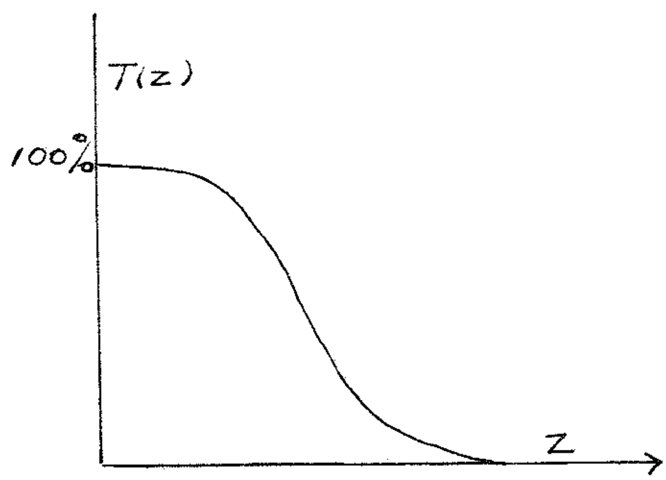
Selection at the Cutoff z:

The number of blocks selected at this cutoff can be expressed in terms of the indicator function for blocks. It is

$$\Sigma \, 1_{Z(vi) \geq z}$$

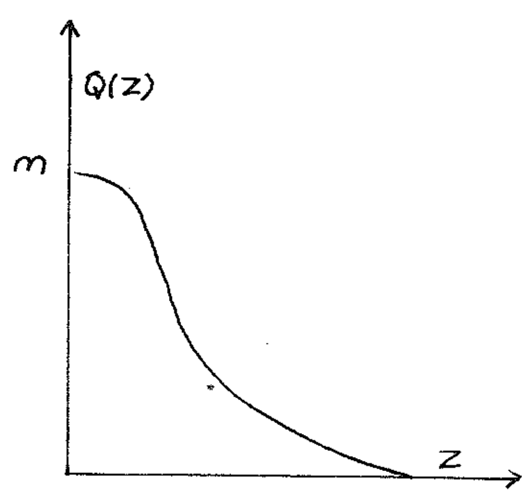
Similarly the tonnage selected is just the proportion of blocks above the cutoff, up to a multiplicative factor T_0. So it is

$$T(z) = 1/n \, \Sigma \, 1_{Z(vi) \geq z} \qquad (14)$$



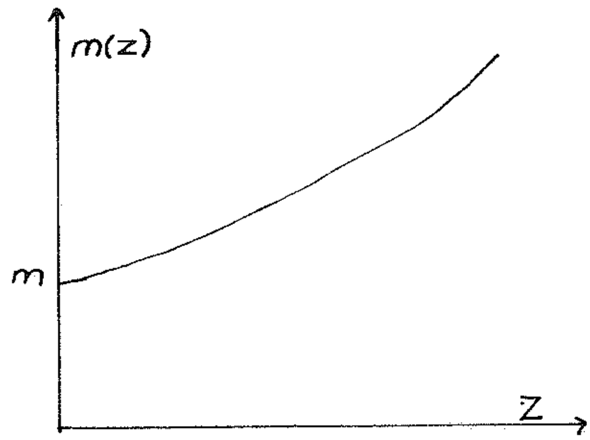
The selected metal is the sum of the metal contained in these blocks. That is,

$$Q(z) = 1/n \, \Sigma \, Z(v_i) 1_{Z(vi) \geq z} \qquad (15)$$



Hence the selected grade is

$$m(z) = \frac{Q(z)}{T(z)} = \frac{\Sigma\ Z(v_i)\ 1_{Z(vi)\geq z}}{\Sigma\ 1_{Z(vi)\geq z}} \qquad (16)$$

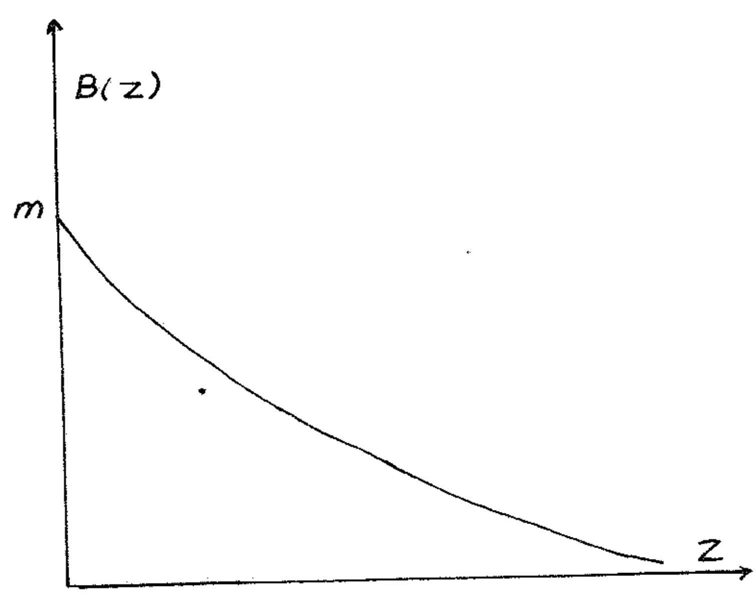


For convenience, we define a conventional profit function:

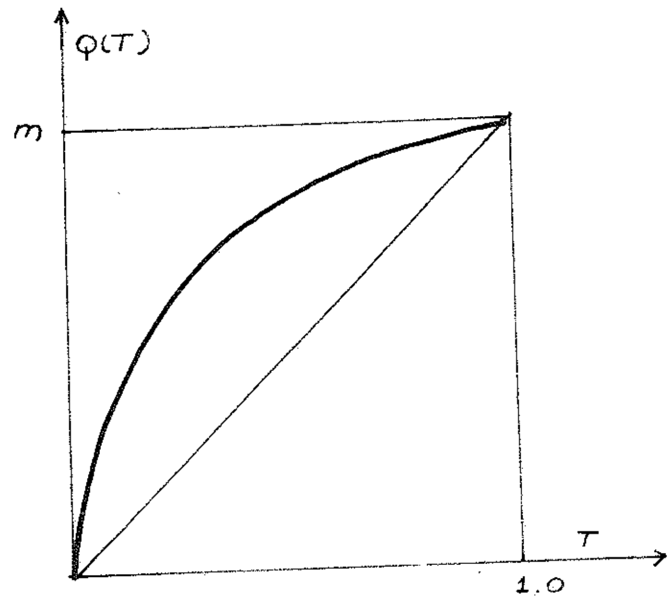
$$B(z) = Q(z) - zT(z) \qquad (17)$$

$$= \frac{\Sigma\ (Z(v_i) - z)\ 1_{Z(vi)\geq z}}{n}$$

Here z is the grade above which the metal in the block pays for its extraction and treatment. When these costs increase (relative to the metal price) z increases and so B(z) decreases. It can be proved that the curve B(z) is convex.



Up to now we have expressed th curves Q, T, Q/T and B as functions of z. But they can also be rewritten as functions of T. For example the curve Q(T) increases (i.e. the quantity of metal recovered increases with increasing ore tonnage) and is concave (i.e. the additional tons selected are poorer and poorer).



<u>Remark</u> Let F(z) denote the cumulative distribution function of the block grades; then $F(z) = P(Z(v) < z)$. It is easy to see that

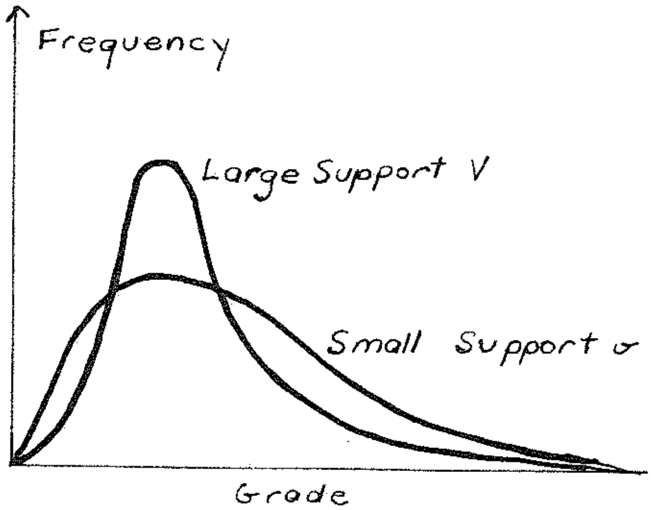
$$T(z) = E(1_{Z(v)\geq z}) = \int 1_{Z\geq z}\, F(dz) = 1 - F(z) \qquad (18)$$

$$Q(z) = E(Z(v)1_{Z(v)\geq z}) = \int Z1_{Z\geq z}\, F(dz) \qquad (19)$$

$$m(z) = E(Z(v)|Z(v)\geq z) \qquad (20)$$

2. SUPPORT EFFECT

How do these curves vary with changing support? We now divide the deposit into big blocks V_i, each with grade $Z(V_i)$. These big blocks are subdivided into smaller ones v. Clearly big blocks are less often very rich or very poor. So their grades are less dispersed than the smaller ones. See Figure.



For a given cutoff z, we have

- for the small blocks: $T_v(z) = 1 - F_v(z)$, $Q_v(z)$, $m_v(z)$ and $B_v(z)$.

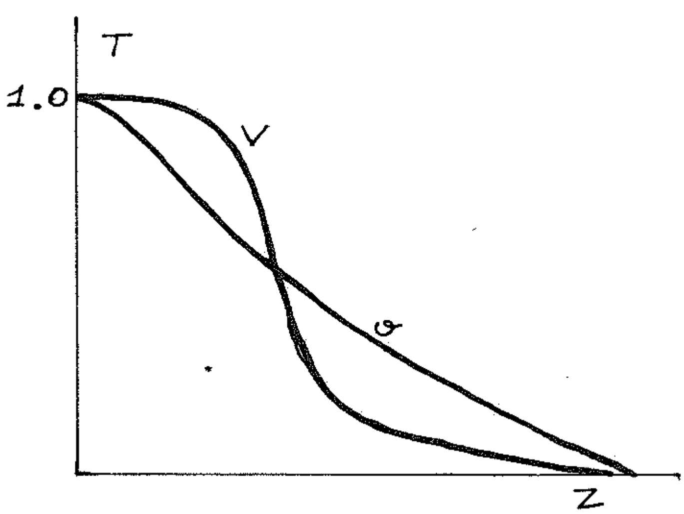
- for the big blocks: $T_V(z) = 1 - F_V(z)$, $Q_V(z)$, $m_V(z)$ and $B_V(z)$.

As their grades are more dispersed, the small blocks generally give a higher mean grade after selection than the big ones:

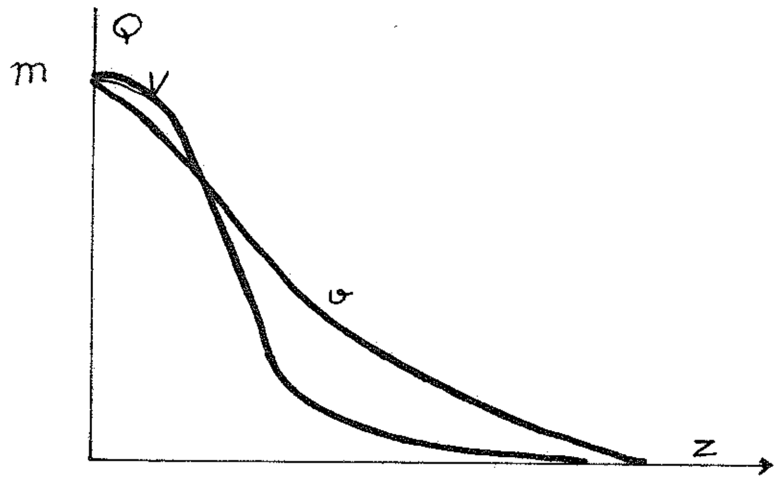
$$m_v(z) > m_V(z)$$

and conversely the remaining small blocks have a lower mean grade than the remaining big ones.

For high cutoffs the proportion of selected blocks and the selected metal tends to be higher for small blocks. Similarly for low cutoffs the proportion of blocks left in place is generally higher for small blocks.

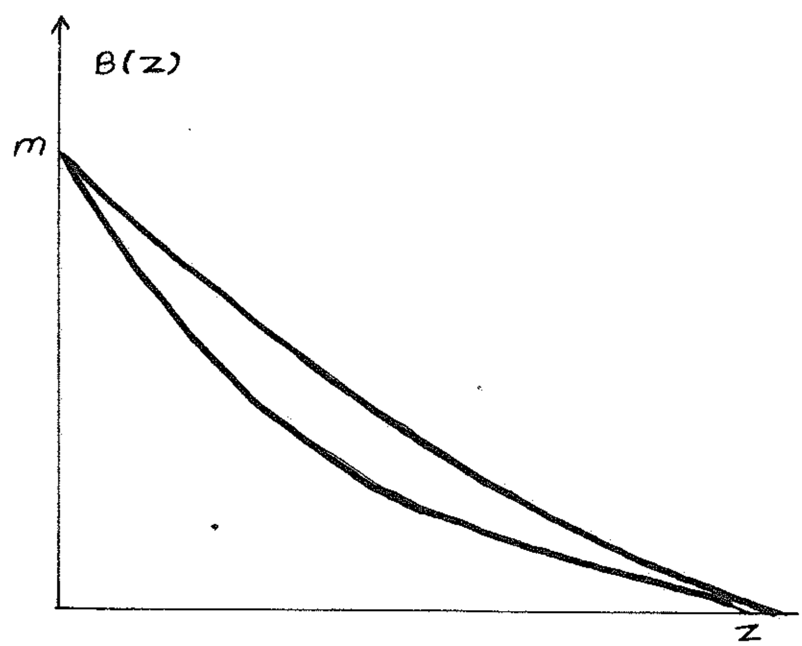


Although the small blocks left in place have a poorer grade than the remaining big blocks, they often represent a larger quantity of metal.

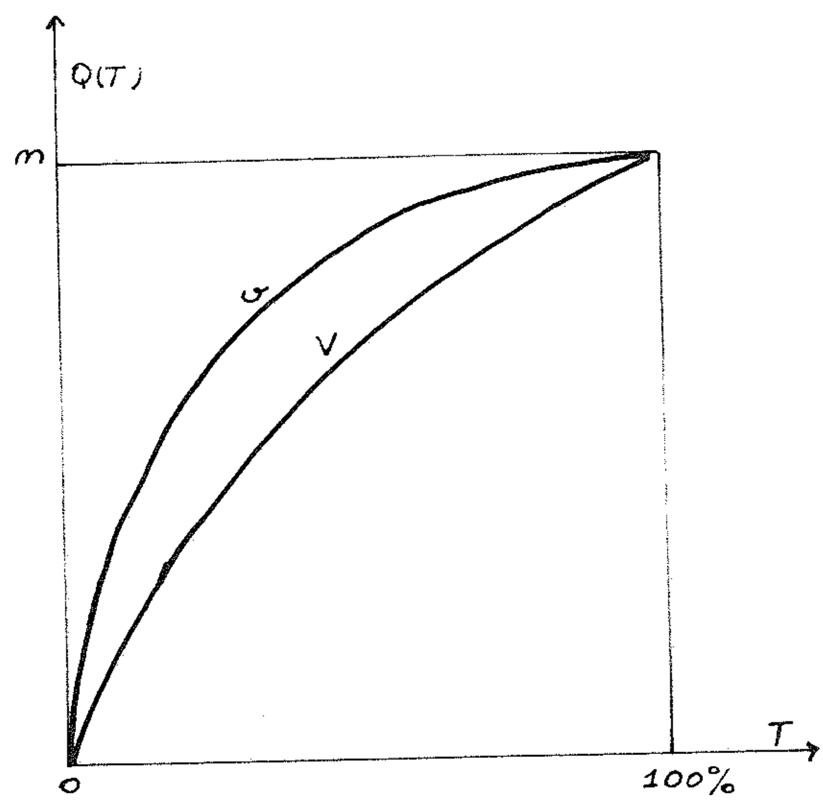


The Conventional Profit Function B(z)

It can be proved that $B_v(z)$ is greater than $B_V(z)$ for any z.



Instead of using the same cutoff for both large and small blocks, we fix the total ore tonnage T to be selected (say to take the richest 50% of the deposit). Then the small blocks will necessarily be richer than the big ones - both in metal and in grade.



3.THE INFORMATION EFFECT

Let us suppose that the deposit is still divided into the big blocks V_i but their grades $Z(V_i)$ are not known. We only have estimates of these, $z^*(v_i)$. Assume that these are conditionally unbiased:

$$E(Z(V)|Z^*(V)) = Z^*(V) \qquad (21)$$

We now can be sure that the grade of the blocks selected as being above the cutoff will be equal to the predicted grade. Conversely, if the conditional bias is not small, the cutoff grade tends to be illusory. For high cutoffs the grade of the selected blocks may even turn out to be below the cutoff!

As $E(Z(V)|Z^*(V)) = Z^*(V)$ and $Z(V) - Z^*(V) = Z(V) - E(Z(V)|Z^*(V))$ are uncorrelated, it follows that

$$Var(Z(V)) = Var(Z^*(V)) + Var(Z(V) - Z^*(V)) \qquad (22)$$

Consequently the estimated grades are necessarily less dispersed (i.e. smoother) than the real ones. (A non-smoothing estimator is always conditionally biased). So these estimated grades are less often very rich or very poorer than are the real ones. Applying a cutoff to them leads a selected grade which is generally smaller, and so is the conventional profit function. This is to be expected. As the selection is made on big blocks, the profit is maximal with perfect information (i.e. when the real values are known).

When we try to select the richest T% of the blocks without their actual grades, some low grade blocks are selected in place of higher grade ones. This results in a decrease in the metal Q(T) and the selected grade Q(T)/T.

4.TOWARDS RECOVERABLE RESERVES:

The simplest case of a selection is when the deposit is divided into small blocks each of size v and the selection is made by comparing the cutoff grade either with (1) the actual block grade (this is the ideal case of perfect information), or (2) its estimated grade obtained from the most detailed information that will be available (e.g. blast hole grades).

In this case the selection is free (from constraints) in the sense that each block is selected or rejected independently of its location and of neighbouring blocks. In principle, knowing the distributions of the real and the estimated grades would allow us to deduce the ore tonnage, the metal content, the grade and the conventional profit function - in other words, the recoverable reserves. But how can we know the distributions at the exploration stage when we only have sample grades at our disposal?

Firstly, these grades are measured on a very small support. Their distribution has not the same characteristics as the selection blocks. It is generally more "selective" in the sense that it suggests better results for the possible selection than is possible for the blocks.

Secondly, these samples are representative of the drillholes, galleries, etc but not the whole deposit. Their experimental distribution might be signicantly different from the true distribution of the samples that would compose the whole deposit. In particular, the experimental sample mean will not be equal to the mean for the whole deposit. It would be interesting to have a confidence interval to see how different is it likely to be.

Let us assume that the samples are punctual (i.e. point sample) and that their distribution is known. Let $F(z)$ denote the cumulative distribution function and let m be the mean and $\gamma(h)$ the variogram. Modelling the distribution $F_v(z)$ of the selection block grades $Z(v)$ would give us the recoverable reserves for the case where the information is perfect. But we have only the mean m and the variance of this distribution: $Var(Z(v) = Var(Z(x) - \gamma(v,v)$. So a change of support model will be required. See the following section for this.

In addition to this, we still have to take acount of the information effect which requires knowing the conditional bias and the distribution of the final estimator.

CHANGE OF SUPPORT MODELS

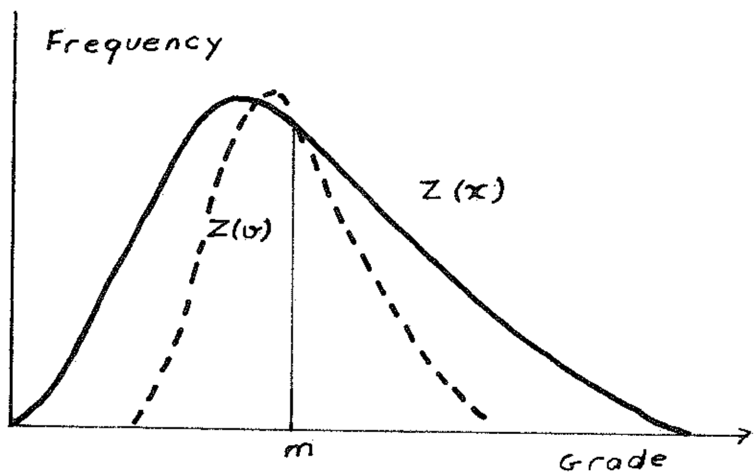
1. AFFINE CORRECTION

In this we obtain an estimate of the selection block distribution by "shrinking" the sample distribution around its mean until it has the correct variance. To be more precise, let σ^2_v be the variance of the selection blocks. Then we have

$$\sigma^2_v = \text{Var}(Z(v)) = \text{Var}(Z(x)) - \bar{\gamma}(v,v)$$

$$= \sigma^2 - \bar{\gamma}(v,v) \qquad (23)$$

The basic assumption is that the variable $(Z(v) - m)$ has the same distribution as $\sigma_v/\sigma\ (Z(x) - m)$. The affinity ratio is then σ_v/σ.



Selecting the richest T% of the deposit would give the recoverable grade:

$m(T)$ for the samples

$m_v(T)$ for the selection blocks

where

$$m_v(T) - m = \sigma_v/\sigma\ (m(T) - m) \qquad (24)$$

For the metal,

$$Q_v(T) = \sigma_v/\sigma\ Q(T) + (1 - \sigma_v/\sigma)mT \qquad (25)$$

It is a very simple model but rather an imprecise one.

2 REVIEW OF STATISTICAL CONCEPTS

(i) Normal Distribution

Let X be a standard normal variate. Its probability density function g(x) is

$$g(x) = 1/\sqrt{2\Pi}\,\exp(-x^2/2) \tag{26}$$

Its cumulative distribution function G(x) is

$$G(x) = \int_{-\infty}^{x} g(t)dt = \Pr\,(X < x) \tag{27}$$

It can be proved that if X is a standard normal variate,

$$E(\exp(\lambda X)) = \exp(-\lambda^2/2) \tag{28}$$

Let Y be normally distributed with mean μ and variance σ^2. Then the random variable

$$X = (Y - \mu)/\sigma \tag{29}$$

is a standard normal variate and we can write

$$Y = \mu + \sigma X \tag{30}$$

(ii) Bivariate Normal Distributions

Let the random variables X and Y have a bivariate normal distribution with a coefficient of correlation ρ. Then any linear combination of X and Y is also normally distributed. Moreover X and Y are independent iff $\rho = 0$.

We now assume that X and Y are standard normal variates (mean 0, variance 1). Then so is the random variable U

$$U = (Y - \rho X)/\sqrt{(1 - \rho^2)} \tag{31}$$

Moreover it is uncorrelated with X and hence independent. Consequently the conditional distribution of Y given X = x can be written as

$$(Y|X = x) = \rho x + \sqrt{(1 - \rho^2)}\;U \tag{32}$$

and so $E(Y|X) = \rho X$ (33)

(iii) Lognormal Variables

Suppose that Z is lognormally distributed; then ln Z is normally distributed with mean μ and variance σ^2. So we can write

$$\ln Z = \mu + \sigma Y \tag{34}$$

where Y is a standard normal variate. The mean and variance of the distribution are

$$m = E(Z) = \exp(\mu + \sigma^2/2) \tag{35}$$

$$\mathrm{Var}(Z) = m^2\,[\exp(\sigma^2) - 1] \tag{36}$$

One particularly convenient way of writing this variable is

$$Z = m\,[\exp(\sigma Y - \sigma^2/2)] \tag{37}$$

The Influence of a Selection

If we write the cutoff grade z_c as

$$z_c = m\,\exp(\sigma y_c - \sigma^2/2) \tag{38}$$

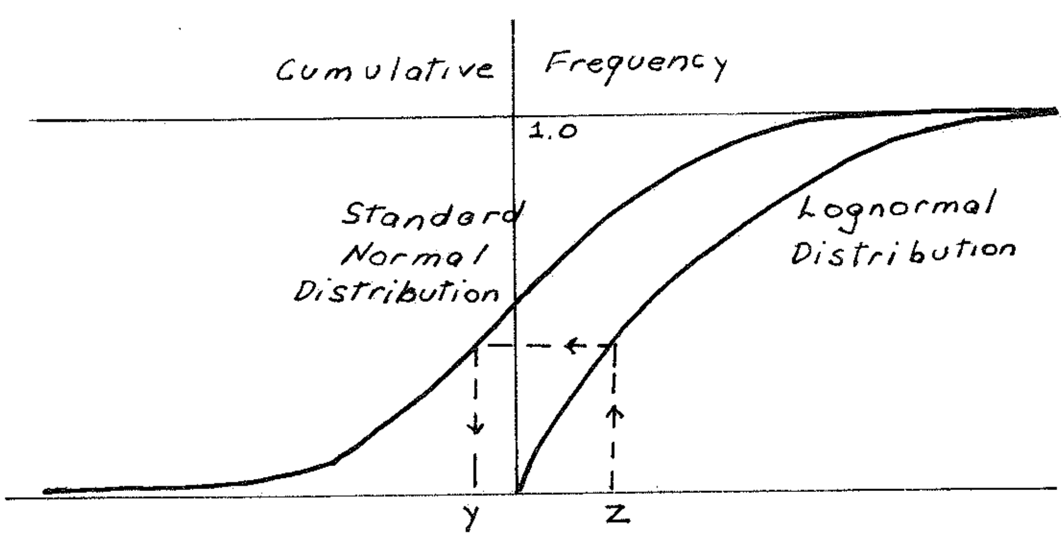
then $Z \geqslant z_c$ implies $Y \geqslant y_c$ and conversely. We obtain the following expressions for the recovery functions.

$$T = E(1_{Z \geqslant zc}) = \Pr(Z \geqslant z_c) = \Pr(Y \geqslant y_c) = 1 - G(y_c) \tag{39}$$

On integrating

$$Q = E(Z1_{Z \geqslant zc}) = m[1 - G(y_c - \sigma)] \tag{40}$$

Remark on the lognormal transformation: The corresponding normal and lognormal variates, Y and Z respectively have the same value of the cumulative distribution function.



3. THE LOGNORMAL MODEL FOR CHANGE OF SUPPORT

(i) Samples: let Z(x) and Y(x) be the lognormal and standard normal variables for the samples.

(ii) Selection Blocks: Suppose that the block grades Z(v) are lognormally distributed. Then

$$Z(v) = m \exp[\sigma_v Y_v - \sigma^2_v/2] \quad (41)$$

where m is the mean of both the sample and block distributions, and where the logarithmic variance is

$$\sigma^2_v = \ln[1 + \mathrm{var}(Z(v))/m^2] \quad (42)$$

The variance of Z(v) is computed from the formula

$$\mathrm{Var}(Z(v)) = \mathrm{Var}(Z(x)) - \gamma(v,v) \quad (43)$$

(iii) Recoverable Reserves

From Eq 38 we can see that

$$y_c = [\sigma^2_v/2 + \ln(z_c/m)]/\sigma_v \quad (44)$$

Hence

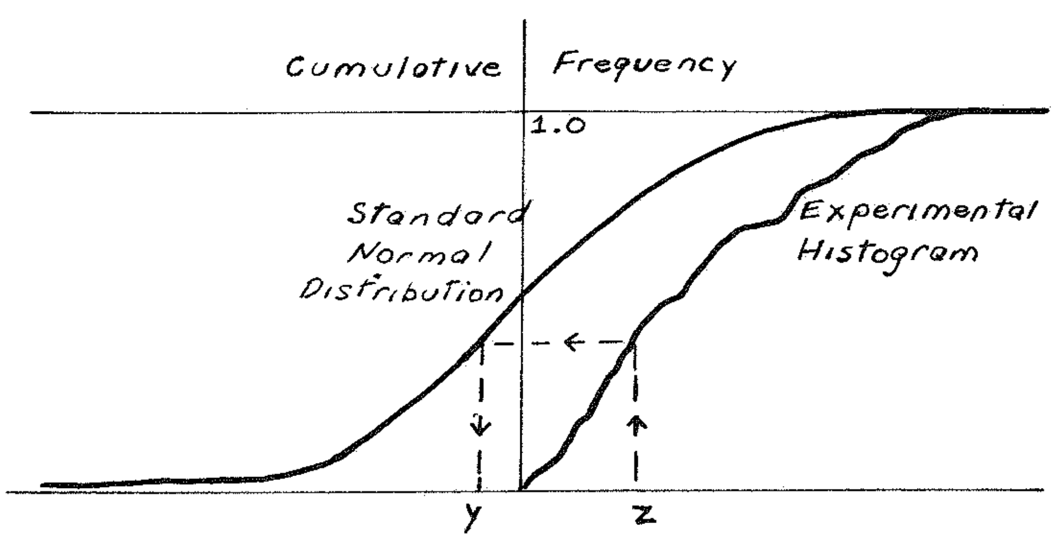
$$T = 1 - G(y_c) \quad (45)$$

$$Q = m(1 - G(y_c - \sigma)) \quad (46)$$

4 GAUSSIAN DISCRETIZED MODEL

(i) The Gaussian Anamorphosis for the Samples

This is just a generalization of the lognormal transformation. We no longer assume that Z is lognormally distributed but merely that there is not a big spike of identical values (e.g. at the origin.)



The experimental histogram of the sample values z is known. Each of these z values can be associated with a value of the standard normal variate y (that is we can write $z = \emptyset(y)$) in such a way that

$$\Pr(Z < z) = \Pr(Y < y) \tag{47}$$

So if we know the cumulative distribution function of Z, then the anamorphosis function $\emptyset$ is defined and conversely.

Given the anamorphosis function $\emptyset$ it is easy to calculate the tonnage and the metal content above a certain cutoff:

$$T = \Pr(Z \geq z_c) = \Pr(Y \geq y_c) = 1 - G(y_c) \tag{48}$$

$$Q = E(Z1_{Z \geq zc}) = \int_{y_c} \emptyset(y)g(y)\,dy \tag{49}$$

(ii) <u>Change of Support</u>

The distribution of Z(v) will be defined via its gaussian anamorphosis function:

$$Z(v) = \emptyset_v(Y_v) \tag{50}$$

where Y_v is as usual a standard normal variate.

How can we find out what the function $\emptyset_v$ is? If $\underline{x}$ is randomly and uniformly located anywhere inside the block v, then

$$E(Z(\underline{x})|Z(v)) = Z(v) \tag{51}$$

Here

$$E[\emptyset(Y_{\underline{x}})|\emptyset_v(Y_v)] = \emptyset_v(y_v) \tag{52}$$

or

$$\emptyset_v(Y_v) = E(\emptyset(Y_{\underline{x}})|Y_v) \tag{53}$$

Now $\emptyset_v$ is fully determined if we know the bivariate distribution of $(Y_{\underline{x}}, Y_v)$. We now make the hypothesis that this is bivariate gaussian. If the correlation coefficient is r, then

$$(Y_{\underline{x}}|Y_v = y) = ry + \sqrt{(1 - r^2)}\,U \tag{54}$$

where U is a standard normal variate. So we have

$$\emptyset_v(y) = \int \emptyset[ry + u\sqrt{(1 - r^2)}]g(u)du \tag{55}$$

This can be written as $\emptyset_r(y)$.

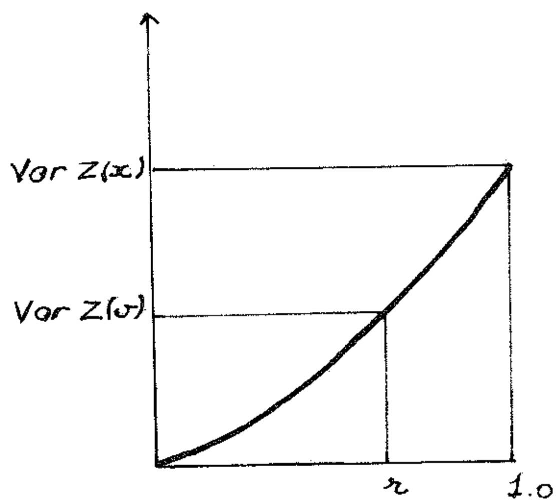
In this model r characterizes the support v:

$$Var(Z(\underline{x})) = Var[E(Z(\underline{x})|Z(v))] + E(Var(Z(\underline{x}))|Z(v))$$

$$= Var(Z(v)) + E(Var(\emptyset(Y_{\underline{x}})|Y_v) \qquad (56)$$

The lower the correlation between $Y_{\underline{x}}$ and Y_v, the larger is $E(Var(\emptyset(Y_{\underline{x}})|Y_v)$ and the smaller $Var(Z(v)) = Var[\emptyset_r(Y_v)]$.

In practice r is calculated to respect the known value of Var(Z(v)).



Remarks: The lognormal model is just one particular case of the gaussian discrete model since if $\emptyset$ is exponential, so is $\emptyset_r$. Secondly, as r decreases toward 0 (i.e. for larger blocks with a small dispersion variance), Z(v) converges toward a gaussian (normally distributed) variable. This is in agreement with the law of large numbers.

ON ANAMORPHOSIS TECHNIQUES

The term "anamorphosis" refers to transformations carried out on the grades (e.g. to make them normally distributed). It can be very helpful during structural analysis and/or when making estimates. But although anamorphosing the grades is often useful, it can sometimes be risky.

1 THE GRADE VARIOGRAM IS NOT ALWAYS ENOUGH TO DESCRIBE THE STRUCTURE

What is Structure?

Basically structure is the set of special spatial relationships that exist between grades Z(x). The most common geostatistical tool used for studying structure is the grade variogram, which is defined as

$$\gamma(h) = 0.5 \text{ Var}[Z(x + h) - Z(x)] \qquad (57)$$

For the distance h the experimental or raw variogram is computed from all the pairs of grades (Z(x+h),Z(x)) which are h apart. Although it is extremely useful, it may turn out to be too poor for describing some phenomena or for use in types of estimation.

Example

Deposits made up of distinct mineralized formations surrounded by waste, where the mining method will follow the outlines of the mineralized zones even though these are unknown at the time when the estimates must be made. In these cases it may be helpful to use indicator functions to describe the geometry of the mineralization:

$$1_{Z(x)>0} = 1 \text{ if } Z(x) \text{ is mineralized}$$

$$= 0 \text{ otherwise} \qquad (58)$$

This also allows us to estimate the proportion of the volume that is mineralized. Although this variable does not describe all the variability found in the grades, it obviously characterizes an important aspect of it.

2. INDICATOR FUNCTIONS AND FACTORS: FITTING BIVARIATE DISTRIBUTIONS

Using a single indicator function when studying deposits with no clearly defined geological cutoff is clearly unsatisfactory. However the cross-structures between all possible indicator functions (i.e. the direct and cross-covariances) provide a rich source of information, which is, in fact, exactly equivalent to knowing the bivariate distributions of (Z(x),Z(x+h)). This can be used to estimate the grades themselves, or to estimate functions of them. In particular, the local recoverable reserves (ore and metal) can be estimated if a suitable change of support formula is available. The problem is to fit a coherent model to the experimental bivariate distribution (in the same way that we must use only authorized variogram models to fit experimental variograms).

The estimation method using indicator cokriging which is called disjunctive kriging, is considerably simplified when an isofactorial model can be fitted to the bivariate distributions. The problem reduces to kriging each of the factors separately. These factors which are functions of the grades, correspond to a rich modelization of the structure which is needed for the subsequent estimation.

3. MULTILOGNORMAL DISTRIBUTION

Another example of an anamorphosis is the lognormal transformation when it converts the raw grades into a normal distribution. This makes it possible to use estimation methods like lognormal kriging or other lognormal estimators. Theoretically these are powerful methods but they are based on the very strong assumption that the distribution is multivariate normal after transformation. As it is difficult to test this assumption, these methods can only be judged a posteriori by the results that they produce.

4. IMPROVING OUR KNOWLEDGE OF THE STRUCTURE

The assumptions in the preceding approach imply that the raw variogram is less structured than that of the transformed grades. This can often be observed in practice. It shows another possible application for anamorphosis: finding a more continuous variogram which is better known than the raw one and so can be taken as being more reliable than it for understanding the structure (e.g. the anisotropies).

Note that variograms are obtained from the bivariate distributions. Consequently if the raw variogram is poorly known then so are the bivariate distributions. This means that hypotheses concerning these laws remain just that - untestable hypotheses.

5. LIMITS AND DANGERS OF THIS IMPROVEMENT

An equally unverifiable hypothesis would then be required to estimate the raw variogram model from that for the anamorphosed data which could be:

- logarithms
- translated logarithms (which reduces the differences between large values without unduly increasing the dispersion among small ones. This has not the same purpose as fitting a three parameter lognormal to obtain a normal distribution)
- indicator functions for a suitable geological cutoff
- truncated grades
- cutting the high values

When using linear estimators of the grades, it would be risky to use the weighting factors obtained from the kriging of more continuous anamorphosed data. The resulting smoothing would be insufficient and it would give each sample too big a weight within its neighbourhood. This would lead to overestimating the areas around rich samples, and consequently, to a dangerous bias if used for selecting mining blocks.

CONFIDENCE INTERVALS ON THE GLOBAL MEAN GRADE

(after J. Rivoirard and C. Lajaunie)

When dealing with skewly distributed grades, the mean grade of samples on a grid may turn out to be different from the true mean grade of the deposit. Considering how uncertain a single estimated value is, it is often more meaningful to have a confidence interval for the mean grade. This is a rather difficult problem. We will only treat it briefly here.

1. SICHEL'S METHOD

Firstly we mention the method proposed by Sichel under hypothesis of <u>independence</u> and <u>lognormality</u>. An estimator for the mean grade and the associated confidence interval are given in Sichel (1966, the March issue of SAIMM).

2. SETTING THE PROBLEM

A fundamental problem in finding a confidence interval for the mean is in defining it. This can best be seen from an illustration. We will consider a deposit V, divided into n equal panels. Please note that we will NOT use the usual geostatistical formalism of a random function, but we will consider that the grade z(x) at a fixed point x is fixed and not random (as in the third part of 'Estimer et choisir", Matheron 1978). Let m be the average grade of this deposit:

$$m = z(V) = 1/V \int_V z(x)\, dx \qquad (59)$$

In one of the panels, we choose a point and we subsequently take this to be the origin of a regular grid whose size is set equal to that of the panels. See the figure below.

Let w denote the support consisting of these n regularly spaced points. Clearly the mean grade of w is just the mean of the n point values:

$$z(w) = \Sigma\, z(x_i)/n \qquad (60)$$

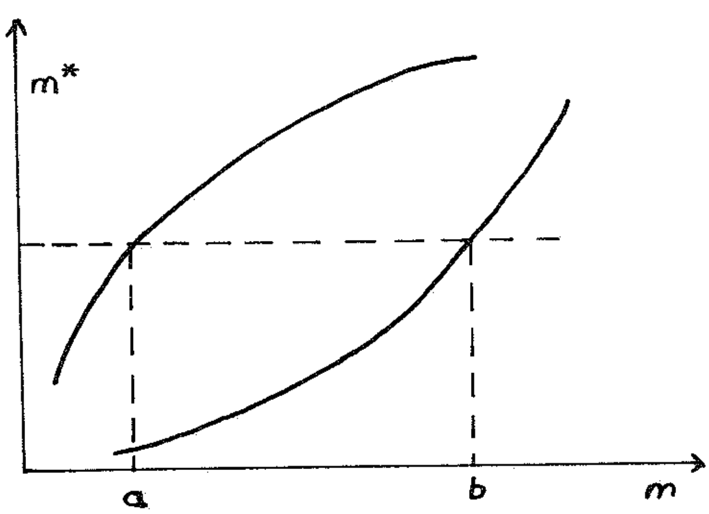
Let us now consider the origin of the grid as being randomly located (with a uniform distribution) within its panel. Let $\underline{w}$ be the randomized support w. Then $z(\underline{w})$ is a random variable that takes a different value each time the origin of the grid is changed. It can be taken as the estimator m* of m. If the distribution of m* was known, then we could easily obtain probability intervals for the variable $z(\underline{w})$. For example, if $z(\underline{w})$ were lognormally distributed with mean m and logarithmic variance $\sigma_w{}^2$, then the 95% probability interval for $z(\underline{w})$ would be given by:

$$\Pr(m \exp(-1.96\, \sigma_w - \sigma_w{}^2/2) \leq z(\underline{w}) < m \exp(1.96\, \sigma_w - \sigma_w{}^2/2)) = 0.95 \qquad (61)$$

But this is <u>not</u> what we are looking for. We want a confidence interval for m. At this stage two approaches for confidence intervals can be taken from the statistical literature.

3. A FIRST APPROACH

Let us suppose that we know the distribution of m* not just for the one real value of m, but for all "possible" values. We could then compute probability intervals for $z(\underline{w})$ for each value of m. See the figure below.



In this first approach the confidence interval for the mean m is obtained by reading the diagram horizontally:]a,b]. This confidence interval is then defined as the set of all values m such that m* belongs to the probability interval associated with m. (Of course since m is not a random variable this interval is not a probability interval for m given the value of m*).

If we suppose that, for all possible m, m* is lognormally distributed, with a logarithmic variance σ_w^2 <u>independent from m</u>, then we have:

$$m^* = a \exp(1.96\ \sigma_w - \sigma_w^2/2)) \qquad (62)$$

$$m^* = b \exp(-1.96\ \sigma_w - \sigma_w^2/2) \qquad (63)$$

and, for the 95% confidence interval]a,b] is

$$]m^* \exp(-1.96\ \sigma_w + \sigma_w^2/2),\ m^* \exp(1.96\ \sigma_w + \sigma_w^2/2)] \qquad (64)$$

There is still the problem of obtaining σ_w^2 in practice. But, in this approach, we are making an hypothesis (σ_w^2 independent from m) that cannot be tested, even if we suppose we know every point grade throughout the deposit. So we go on to the second approach.

2. SECOND APPROACH

In this approach (see for instance: Rohatgi, "An introduction to probability theory and mathematical statistics", Wiley, 1976), there is no need to imagine different values for m, which is conceptually more satisfactory. Briefly a confidence interval is defined as a random interval which is a function of the random variable $z(\underline{w})$ and which contains the true value m with a minimum given probability.

Theoretically, it is possible to verify this result by taking all possible origins of grid. consequently this type of confidence interval can be said to have an objective significance. Moreover the hypothesis that the confidence interval is built on can theoretically be tested in the same way.

Before going into the hypotheses to be made, let us look at the information that is (or can be hoped to be) available in practice from a single grid w of samples:

- the histogram of these point values

- their mean value z(w)

- their variance $\sigma^2_{x|w}$

- their variogram $\gamma_w(h)$

Note: the variance of the samples $\sigma^2_{x|w}$ is equal to the mean variogram over w: $\gamma_w(w,w)$.

This variogram $\gamma_w(h)$ (even if known) cannot reasonably be considered as equal to the unknown point variogram on V i.e. $\gamma(h)$. But it is often reasonable to <u>suppose</u> (the first hypothesis!) that it only differs fom it by a multiplicative factor. As $\gamma(h)$ is theoretically the mean value of $\gamma_w(h)$ when w is randomized (at least for h that are multiples of the lag), it gives

$$\frac{\gamma(h)}{\gamma_W(h)} = \frac{E(\sigma^2{}_{x|w})}{\sigma^2{}_{x|w}} \qquad (65)$$

We know that we can calculate variances once the variogram is known. Let $\underline{w}$ be a random grid and let x be a point in $\underline{w}$ chosen at random; then

$$E(z(x)|\underline{w}) = z(\underline{w}) \qquad (66)$$

On writing

$$\mathrm{Var}(z(x)) = \mathrm{Var}[E(z(x)|\underline{w})] + E[\mathrm{Var}(z(x)|\underline{w})] \qquad (67)$$

where $\mathrm{Var}(z(x)) = \gamma(v,v)$ (68)

and $E[\mathrm{Var}(z(x)|w)] = E(\sigma^2{}_{x|w}) = E[\gamma_W(w,w)] = \gamma(w,w)$ (69)

we obtain

$$\mathrm{Var}(z(\underline{w})) = \gamma(v,v) - \gamma(w,w) \qquad (70)$$

From (65) and (70) we deduce:

$$\frac{\mathrm{Var}(z(\underline{w}))\ \sigma^2{}_{x|w}}{E[\sigma^2{}_{x|w}]} = \gamma_W(V,V) - \gamma_W(w,w) = \text{constant} \qquad (71)$$

Here we have to make the crucial hypothesis, and we will choose the lognormal case. We assume that the distributions are approximately lognormal; that is,

$z(x)|w$ (sample values for each grid w) is lognormal with mean z(w) and with logarithmic variance σ^2 independent of w,

$z(\underline{w})$ is lognormal with mean m and with logarithmic variance $\sigma_w{}^2$

Of course this implies that the experimental distribution of samples can already be considered as lognormal (the other hypothesis being not testable unless every point grade of the deposit is known).

Consequently, we can write

$$z(\underline{w}) = m \exp[\sigma_w y_w - \sigma_w{}^2/2] \qquad (72)$$

where y_w is a standard normal variate. Then it follows that

$$\mathrm{Var}(z(\underline{w})) = m^2\,[\exp(\sigma_w^2) - 1] \tag{73}$$

$$E(z(\underline{w})^2) = m^2 \exp(\sigma_w^2) \tag{74}$$

$$\sigma^2_{x|w} = z_w^2\,[\exp(\sigma^2) - 1] \tag{75}$$

$$E(\sigma^2_{x|\underline{w}}) = m^2 \exp(\sigma_w^2)\,[\exp(\sigma^2) - 1] \tag{76}$$

Thus

$$\frac{\mathrm{Var}(z(\underline{w}))\sigma^2_{x|w}}{E(\sigma^2_{x|w})} = \frac{[\exp(\sigma_w^2) - 1]z_w^2}{\exp(\sigma_w^2)} = \text{constant} \tag{77}$$

This gives $\sigma_w^2 = -\ln[1 - \text{constant}/z_w^2]$ (78)

Then for instance

$$\Pr(-1.96 \leq y_w < 1.96) = 95\% \tag{79}$$

gives

$$\Pr(m \exp(-1.96\,\sigma_w - \sigma_w^2/2) \leq z_{\underline{w}} < m \exp(1.96\,\sigma_w - \sigma_w^2/2))$$

$$= \Pr(z_{\underline{w}} \exp(-1.96\,\sigma_w + \sigma_w^2/2) < m \leq z_{\underline{w}} \exp(1.96\,\sigma_w + \sigma_w^2/2))$$

$$= 95\% \tag{80}$$

The random interval given below is then a 95% confidence interval:

$$]\; z_{\underline{w}} \exp(-1.96\,\sigma_w + \sigma_w^2/2),\; z_{\underline{w}} \exp(1.96\,\sigma_w + \sigma_w^2/2)\;]$$