

Applied Geostatistics for Reservoir Characterization

1 Chap. 2: Principles of Statistics

- The relative frequency (f_{Ri}) is the number of measurement falling within particular class i (f_i) over the number of data (n) ($f_{Ri} = f_i/n$ with $\sum f_{Ri} = 1$). Cumulative relative frequency (F_j) sum up the frequency for higher class ($F_j = \sum f_{Ri}$).

- Descriptive statistic:

$\bar{x} = 1/n \sum x_i$	mean
$x = x_i, f_i > f_j \forall j$	mode
$\tilde{x} = x_i, x_{u < i} = n/2$	median
$x_p\% = x_i, x_{u < i} = p\% \times n$	percentile
$s^2 = 1/(n-1) \sum (x_i - \bar{x})^2$	Variance
$C_v = s/\bar{x}$	Coef. variation
$R = x_{max} - x_{min}$	Range

- Spacial data sets. In spacial sampling, bias is introduced when preferential localisation of sampling is used. Sample de-clustering method divided data in subarea and statistical property is computed by assigning a weight to each data according to the number of data in its subarea. Moving-Windows Statistics compute statistical summary on a small subset of the data falling within a defined window.

- Bivariate Distributions. Covariance and Correlations:

$$C(x, y) = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum x_i y_i - \bar{x} \bar{y}$$

$$r(x, y) = \frac{C(x, y)}{s_x s_y}, \quad -1 < r < 1$$

- Linear Regression. The simple linear equation ($y = mx + b$) present parameter which can be best fit with $m = C(x, y)/s_x^2$ and $b = \bar{y} - m\bar{x}$.
- Probability. Probability that an event A occurs can be defined as $p(A) = \lim_{n \rightarrow \infty} n_a/n$ (n_a : number of times where A occurs over n times). Basic law of prob. states that $0 \leq p(A) \leq 1$, $p(A) + p(\bar{A}) = 1$ and for two events A and B, $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.

- Conditional probability helps to compute the prob. of an event A knowing about the event B such that:
 $p(A|B) = \frac{p(A \cap B)}{p(B)}$. This can be explained as reducing the sample space to event B.
This idea can be further enlarged to multiple subsets of A which compute together the entire space ($\sum p(A_i) = 1$). So that an event B can be expressed by subset of A:
 $p(B) = \sum p(A_i \cap B)$. This gives the Bayes Theorem:

$$p(A_j|B) = \frac{p(A_j \cap B)}{\sum p(A_i \cap B)} = \frac{p(B|A_j)p(A_j)}{\sum p(B|A_i)p(A_i)}$$

- Probability function. The probability mass function ($P(a)$) of a random variable X is $P(a) = p(X = a)$. And for continuous variable, the probability density function $p(a < X < b) = \int_a^b f(x)dx$. Cumulative distribution function make a cumulative sum: $F(a) = p(X < a) = \int_{-\infty}^a f(x)dx$

- Important Distribution Function.

Name	Uniform	Normal	Log-Normal
$f(x)$	$\frac{1}{b-a}$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$
$F(x)$	$\frac{x-a}{b-a}$	$\frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x-\mu}{\sigma\sqrt{2}} \right) \right]$	$\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left[\frac{\ln x - \mu}{\sqrt{2}\sigma} \right]$
$E(x)$	$\frac{a+b}{2}$	μ	$e^{\mu + \sigma^2/2}$
$V(x)$	$\frac{(b-a)^2}{12}$	σ^2	$(e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$

- Inference allows estimation of parameters from sample statistic. Unbiased stats that the true parameters are close to the estimated parameter such that $E[\hat{\theta}] = \theta$. Minimum variance states that the variance of the distribution of the estimated parameter should be as small as possible. From these two properties, arise the MVUE technique.

2 Chap. 3: Spacial Relationship

2.1 Stationarity

First order: $f(X(\mathbf{u})) = f(X(\mathbf{u} + \mathbf{L}))$ Second order:
 $f(X(\mathbf{u}), X(\mathbf{u} + \mathbf{h})) = f(X(\mathbf{u} + \mathbf{h}), X(\mathbf{u} + \mathbf{h} + \mathbf{L}))$

2.2 Variogram

variogram measure the variance of the difference from two point separate by \mathbf{L} .

$$\gamma(\mathbf{L}) = \frac{1}{2} \operatorname{Var} [X(\mathbf{u}) - X(\mathbf{u} + \mathbf{L})]$$

Expanding this definition with first-order stationarity arise to a simplified version

$$\gamma(\mathbf{L}) = \frac{1}{2} E [(X(\mathbf{u}) - X(\mathbf{u} + \mathbf{L}))^2] = C(0) - C(\mathbf{L})$$

Empirical variogram is built by setting a lag distance (ΔL) and lag orientation ($\Delta \theta$). The use of Bandwidth helps in point far away.

Name	$x_T(u)$
log-transform	$\log_{e,10}(x(u))$
power-transform	$[x(u)]^p$
Rank	$i/(n+1)$
Indicator transform	1 for $x(u) < x_t$ else 0
Normal-score	Empirical and Normal CDF

- The General-Relative Variogram normalized the conventional variogram with local mean and eliminates the influence of the variation in lag-mean.
- Pairwise-Relative Variogram also normalized but by the pair mean rather than lag mean.
- Non-ergodic Variogram
- The following parameters are often used to describe variogram:
 - nugget n : The height of the jump of the semi-variogram at the discontinuity at the origin.
 - sill s : Limit of the variogram tending to infinity lag distances.
 - range r : The distance in which the difference of the variogram from the sill becomes negligible.

- In modelling variogram, two rules need to be applied (1) uses of minimum number of param. and (2) positive definiteness.

Name	$\gamma(h)$	$C(L)$
Exponential	$C_0 \left[1 - \exp \left(\frac{-3L}{r} \right) \right]$	$C_0 \exp \left(\frac{-3L}{r} \right)$
Spherical	$C_0 \left[\frac{3L}{2r} - \frac{L^3}{2r^3} \right]$	$C_0 \left[1 - \frac{3L}{2r} + \frac{L^3}{2r^3} \right]$
Gaussian	$C_0 \left[1 - \exp \left(\frac{-3L^2}{r^2} \right) \right]$	$C_0 \exp \left(\frac{-3L^2}{r^2} \right)$
no-sill (Fractional Gaussian Noise or Brownian Motion)		
Logarithmic	$C_0 \log L$	
Sin	$C_0 \left[1 - \frac{\sin(rL)}{L} \right]$	$C_0 \frac{\sin(rL)}{L}$
Cos	$C_0 [1 - \cos(rL)]$	$C_0 \cos(rL)$

- Cross Variogram

$$\gamma(\mathbf{L}) = \frac{1}{2} E [X(\mathbf{u}) - X(\mathbf{u} + \mathbf{L})] [Y(\mathbf{u}) - Y(\mathbf{u} + \mathbf{L})]$$

- Multipoint Histogram allows definition of the connectivity of multipoint instead of only 2 points as the variogram does.

3 Conventional Estimation

- Simple Kriging is using MVUE, resulting in the linear

$$X^*(\mathbf{u}_0) = \lambda_0 + \sum \lambda_i X(\mathbf{u}_i) \begin{cases} E[X^*(\mathbf{u}_0) - X(\mathbf{u}_0)] = 0 \\ \min \sigma^2 [X(\mathbf{u}_0) - X^*(\mathbf{u}_0)] \end{cases}$$

system $|\mathbf{A}| = |\mathbf{C}|^{-1}|\mathbf{c}|$ can be solved and the error variance of the estimate: $\sigma_E^2 = C(u_0, u_0) - \sum \lambda_i C(u_i, u_0)$. The weight (λ) is assigned according to his spatial relationship to (1) the unsampled location and (2) to the other point. The error variance is equal to variance in the absence of spacial information. Variogram could replace covariance in all equations. Simple kriging assumes constant mean over space.

- Cross validation

- Block Estimation takes into account that the estimated location is not a point but as an area associate. It does the same thing as compute point and arithmetic average (not good for permeability) then within the cell but does it more efficiently.

- Ordinary Kriging doesn't assume first-order stationarity (mean). It uses the method of the Lagrangian multiplier and finishes in a similar matrix.

- Cokriging is used to estimate one variable (primary) value knowing other sampled variables (covariables). It uses the assumption that the variables are spatially related with a linear relationship. The relationship must be strong: (1) high correlation, (2) physical link and (3) good history matching. This procedure is very computer expensive.

- Universal kriging procedure is used in presence of a trend.

- Log-Normal Kriging is a non-linear kriging procedure which (1) reduces the variability of sample point and (2) transform log-normal distributed variables in normal distribution. The back-transform depends on error variance. (not very used)

- *Multi-gaussian* convert with a normal z-score the data.
- *Indicator Kriging* final result represent the uncertain distribution at unsampled location (instead of single value) and has not simple back transform which therefore require a conditional simulation technique. Indicator variable express the confidence in the sample values. It therefore allows incomplete or soft information (value less than 1). There are not distribution assumed in the data.

$$I(u_j, x_t) = \begin{cases} 1, & \text{if } X(u_j) < x_t \\ 0, & \text{if } X(u_j) > x_t \end{cases}$$

4 Conditional Simulation

Three characteristic distinguish conventional estimation to conditional simulation:

1. *Sample Variability*. Because of the linear relationship, conventional estimation never reproduce extreme value while conditional simulation can keep the distribution very well if desire.
2. *Sample Spatial relationship*. Firstly, because of 1., the variance of the estimation is smaller than the observation and therefore the variogram as a smaller sill. Secondly, because estimation does not use previously estimated value in the construction of the weight, some of the spacial relationship can't be kept. Simulation technique reuse simulated value as known value and preserve the variogram better.

3. *Uncertainty*. In conventional estimation, the error variance σ_E describe surrounding error but not local. Conditional simulation build equal likely realisation. Uncertainty can be estimate from these different possible realisation.

5 Grid-base Simulation Methods

5.1 Sequential

5 steps:

1. Transform (Indicator and Gaussian)
2. Model variogram
3. Select random path
4. Sequentially estimate value. Simulation depend on the path because the previously simulated value are reused. We can set a max nb of value reused or ratio data/simulated value.
5. Back transform

5.2 Cosimulation

Cosimulation extend sequential simulation by simulating multiple attributes simultaneously. It is based on a cokriging approach and require linear relationship between attribute.

Again, the procedure follows the same 5 steps:

1. Transform: Indicator for discrete and Gaussian for continuous. We need linear relationship in transform domain between the attribute in order to apply cokriging technique.

2. Model variogram : one for every threshold value in indicator transform but also cross-variogram for each dependences
3. Select random path
4. Sequentially estimate value. We simulate in order from the least dependent attribute to the most dependant. For cokriging procedure, assuming gaussian distribution:

$$X_k^*(\mathbf{u}_0) = \sum^{n_k} \lambda_{i_k} X_k(\mathbf{u}_i) + \sum^L \sum^{n_l} \lambda_{j_l} X_l(\mathbf{u}_j)$$

5. Back transform

5.3 Probability Field Simulation

Once the posteriory distribution is known, a uniform random number is pick up to determine his value. Probability field simulations minimizes the short-scale variability by creating a spacial correlation in the number pick up

5.4 Simulated Annealing

This method honor univariate statistic, spacial relationship and relationship with other attribute. It futher incorporate other constraints, however, it is very computer expensive.

6 Scaleup