# **Principles of Statistics**

## Introduction

## Descriptive Statistics

Descriptive Statistics deals with understanding of the type of data currently available or the **sample data**.

### Frequency Distribution

### Summary Statistics for Univariate Distribution

#### Mean

#### Median:

Central value after ranking, or the value at 50th percentile.

#### Mode:

(hint: bi-modal, multi-modal): Value occurring at highest frequency

#### Extremes:

#### Percentile:

#### Variance:

For Linear statistics:

#### Standard Deviation:

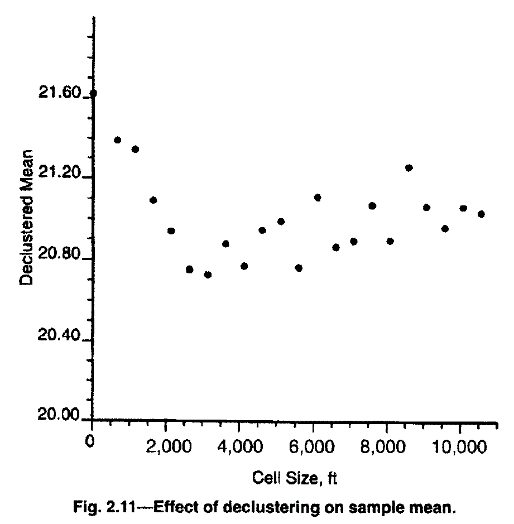
For Linear statistics:

#### Coefficient of variation:

#### Range

### Spatial Data Sets

#### Sample declustering

* + Step 1: Weights for each points are calculated:
  + Step 2: Arithmetic mean of samples after calculating weights:
  + Step 3: What size the subarea should be?
    - Option 1:
    - Option 2: Different subareas should be tried and declustered mean plotted against the subarea size. The size should be chosen where the declustered mean reaches maximum or minimum. E.g. while calculating the porosity usually wells are drilled in the sweet spots, thus there are more sample values of higher porosity. Hence, the subarea where the minimum average porosity is reached is chosen.
  + Since the declustered data is more uniform, it has more variance/standard deviation than the original clustered data.

#### Moving Window Statistics

* + A small window of desired size is chosen and all the samples within that window is used to calculate the local summary statistics e.g. mean, variance, s.d. If *local means and variances* are uniform it is called *homoscedastic* and if there is significant variation it is called *heteroscedastic*.
  + Choosing the right window size and shape: Recommended is a rectangular window and large enough with enough sample points inside, so that two adjacent windows can overlap and have common samples.
  + Several possibilities exist: Local mean and variance can both vary, local mean can stay fairly constant and local variance can vary and vice versa. In many earth science problems, local mean has been observed to be proportional to local standard deviation which is called *proportional effect*.

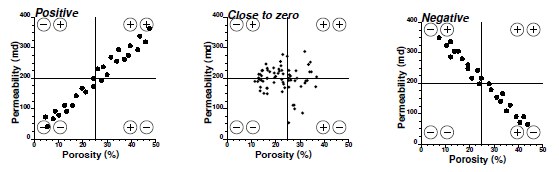
### Bivariate Statistics

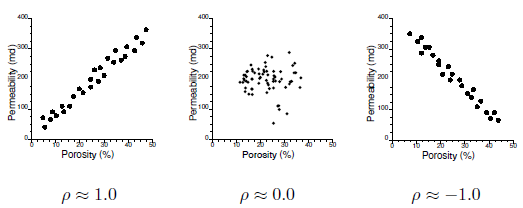
#### Conditional frequency distribution

#### Summary Statistics for Bivariate Distribution

##### Covariance:

For Linear statistics:

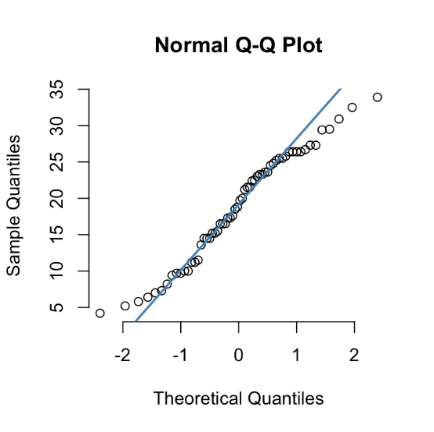




##### Correlation Coefficient

##### Rank correlation coefficient:

Steps: All data are sorted in ascending order and assigned a rank, depending on where it falls where the smallest number gets the lowest rank. With each value assigned a rank, the correlation is calculated for the ranks from the two sets of data, instead of the values themselves.

**Q-Q Plots**: Graphical method of comparing two probability plots by plotting their quantiles against each other.

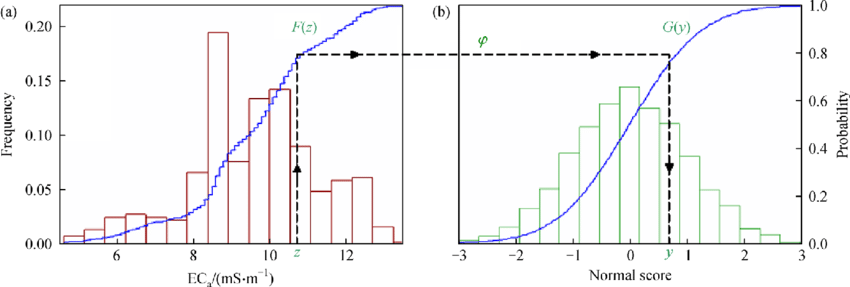
If points lie on x=y line, the two distributions are similar.

If points lie on a line, the two distributions are linearly related.

Often used to compare a dataset against a pre-existing model.

**Normal score transformation**: A given data point is assigned a value which is either exactly, or an approximation, to the expectation of the order statistic of the same rank in a sample of standard normal random variables of the same size as the observed data set. (It preserves the expected value of each data point.)

In the below diagram, point z is transformed to point y. (so for n ‘z’ points, we generate n ‘y’ points)



##### Linear Regression

Establishing relationship between two variables, so as to predict the value of one variable when the value of the other variable is known.

##### Bivariate relationship for spatial data:

See the Chapter 2 on the relationships of spatial data.

## Inferential Statistics

It is a logical extension of the descriptive statistics, where it deals mostly with sample data sets. However, from the characteristics of the samples conclusions can be drawn about the population (from which the sample was taken). Inferential statistics is known to handle this kind of problems.

How ‘samples statistics’ are different from the ‘inferential or population statistics’? Jerry Jensen UPDC notes

What is Random Experiment? (Kelkar Perez)

Conditional probability (Kelkar Perez, Jerry Jensen UPDC notes)

Bayes Theorem (Kelkar Perez, Jerry Jensen UPDC notes)

### Random Experiment

A random function model replaces reality by a set of ‘alternate possible realities’ & all these ‘alternate possible realities’ are realizations of a random function. All these ‘alternate possible realities’ share a few properties i.e. mean, variance and covariance. At any location, there are a series of possible values determined by the random function model and any of these realizations can be the real one with certain probability factor. (So there is always an uncertainty associated with any given number.)

Small case alphabet **z**, denotes simply a ‘deterministic’ variable

Uppercase alphabet **Z**, denotes a Random Variable

A random function is a rule that assigns a realization to the outcome of an experiment.

Each random variable is characterized by a probability distribution function (pdf)

E.g., a Gaussian dist.

[Jamie UPDC notes]

If we are able to develop a completely deterministic model based on evolution of reservoir

Input from Kelkar Perez.

### Sample Space and Events

Sample space is the set of all possible outcomes of a random experiment.

### Probability

#### Laws of probability

#### Conditional Probability

###### Bayes Theorem

### Random Variables

A random variable is a variable whose values are generated by a random experiment on the basis of some probabilistic function.

#### Probability Function

#### Cumulative Distribution Function

#### Bivariate Functions

### Mathematical Expectation

#### Expected Value

##### Characteristics of Expected Value

#### Important parameters of Univariate Distribution

##### Arithmetic mean

##### Variance

For discrete random variable:

Sample standard deviation is denoted by **s** and s.d. of random variable is denoted by **σ**

For continuous random variable:

###### Characteristics of variance:

##### Moment

The n-th moment of a real-valued continuous function f(x) of a real variable about a value c is:

* The first **raw moment** is the mean or [When c=0, is raw moment].
* The 2nd central moment is the variance.

##### Central moment

In probability theory and statistics, a central moment is a moment of a probability distribution of a random variable about the random variable's mean **µ**; that is, it is the expected value of a specified integer power of the deviation of the random variable from the mean.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Significance of moments (raw, central, normalised) and cumulants (raw, normalised), in connection with named properties of distributions | | | | | |
| **Moment  ordinal** | **Moment** | | | **Cumulant** | |
| **Raw** | **Central** | **Standardized**  where | **Raw** | **Normalized** |
| 1 | Mean | 0 | 0 | Mean | N/A |
| 2 | – | Variance | 1 | Variance | 1 |
| 3 | – | – | Skewness | – | Skewness |
| 4 | – | – | (Non-excess or historical) kurtosis | – | Excess kurtosis |
| 5 | – | – | Hyperskewness | – | – |
| 6 | – | – | Hypertailedness | – | – |
| 7+ | – | – | – | – | – |

##### Standardized moment

The standardized moment of degree k is . () The standardized moment is scale invariant.

The first four standardized moments can be written as:

|  |  |  |
| --- | --- | --- |
| Degree *k* |  | Comment |
| 1 | 0μ ~ 1 = μ 1 σ 1 = E ⁡ [ ( X − μ ) 1 ] ( E ⁡ [ ( X − μ ) 2 ] ) 1 / 2 = μ − μ E ⁡ [ ( X − μ ) 2 ] = 0 {\displaystyle {\tilde {\mu }}\_{1}={\frac {\mu \_{1}}{\sigma ^{1}}}={\frac {\operatorname {E} \left[(X-\mu )^{1}\right]}{(\operatorname {E} \left[(X-\mu )^{2}\right])^{1/2}}}={\frac {\mu -\mu }{\sqrt {\operatorname {E} \left[(X-\mu )^{2}\right]}}}=0} | The first standardized moment is zero, because the first moment about the mean is always zero. |
| 2 | μ ~ 2 = μ 2 σ 2 = E ⁡ [ ( X − μ ) 2 ] ( E ⁡ [ ( X − μ ) 2 ] ) 2 / 2 = 1 {\displaystyle {\tilde {\mu }}\_{2}={\frac {\mu \_{2}}{\sigma ^{2}}}={\frac {\operatorname {E} \left[(X-\mu )^{2}\right]}{(\operatorname {E} \left[(X-\mu )^{2}\right])^{2/2}}}=1} 1 | The second standardized moment is one, because the second moment about the mean is equal to the variance σ2. |
| 3 | μ ~ 3 = μ 3 σ 3 = E ⁡ [ ( X − μ ) 3 ] ( E ⁡ [ ( X − μ ) 2 ] ) 3 / 2 {\displaystyle {\tilde {\mu }}\_{3}={\frac {\mu \_{3}}{\sigma ^{3}}}={\frac {\operatorname {E} \left[(X-\mu )^{3}\right]}{(\operatorname {E} \left[(X-\mu )^{2}\right])^{3/2}}}} | The third standardized moment is a measure of skewness. |
| 4 | μ ~ 4 = μ 4 σ 4 = E ⁡ [ ( X − μ ) 4 ] ( E ⁡ [ ( X − μ ) 2 ] ) 4 / 2 {\displaystyle {\tilde {\mu }}\_{4}={\frac {\mu \_{4}}{\sigma ^{4}}}={\frac {\operatorname {E} \left[(X-\mu )^{4}\right]}{(\operatorname {E} \left[(X-\mu )^{2}\right])^{4/2}}}} | The fourth standardized moment refers to the kurtosis. |

But **coefficient of variation** is (measure of the spread of the distribution) , the reciprocal of the first standardized moment.

##### Moment Generating Function

The Moment generating function of a random variable **X** (of pdf **f(x)**) is

always exists and is equal to 1. However, a key problem with moment-generating functions is that moments and the moment-generating function may not exist, as the integrals need not converge absolutely. By contrast, the characteristic function or Fourier transform always exists (because it is the integral of a bounded function on a space of finite measure), and for some purposes may be used instead.

The nth moment is:

***Relation of MGF to other functions:***

1. Characteristic function is the moment generating function of iX or moment generating function of X evaluated on imaginary axis iX.
2. Cumulant generating function
3. Probability generating function:

##### Cumulant

In probability theory and statistics, the cumulants of a probability distribution are a set of quantities that provide an alternative to the moments of the distribution. The moments determine the cumulants in the sense that *any two probability distributions whose moments are identical will have identical cumulants as well, and similarly the cumulants determine the moments*.

can be obtained by differentiating n times and evaluating the result to 0.

If the moment-generating function *does not exist*, the cumulants can be defined in terms of the *relationship between cumulants and moments*.

***Alternative definition of cumulant generating function:***

Some writers prefer to definite the cumulant generating function as the natural logarithm of the **characteristic function** . Hence the is sometimes called the *second characteristic function*.

* Adv: is well defined for all real values of t even when is not well defined for all real values of t.
* Use in statistics:

For statistically independent random variables X and Y.

***Properties of cumulants:***

* Invariance and equivariance:

* Homogeniety
* Additivity

* A negative result

Normal distribution doesn’t have cumulants for m>3 with non-zero lower order cumulants.

#### Important properties of Bivariate Distribution

##### Expectation

##### Covariance

For C[X,Y]=0 means the X & Y variables are independent.

For X=Y, C[X,Y]=V[X]

##### Some properties of covariance

Variance of single variable composed of multiple variables:

Covariance of multiple variables

##### Correlation coefficient

* Population Corr. Coeff.:
* Sample Corr. Coeff.: **r**

### Important Distribution Functions

#### Uniform Distribution

#### Normal (Gaussian) Distribution

#### Log-Normal Distribution

### Inference of Parameters (MVUE)

In terms of variables x:

For random variables, true value:

For random variables, estimated value:

Two requirements are to be applied to the error : unbiasedness & variance minimization

#### Unbiasedness

For unbiasedness condition, the expected value of the error should be 0.

#### Minimum variance

The variance of the error has to be kept minimum.

To minimize the error variance, differentiating w.r.t. and :

Solving for :

Or

Solving for :

#### Generalized linear regression

A generalized expression for estimating Y from random variables.

Applying unbiased condition:

Applying the minimum variance condition:

To minimize variance, we need to equate partial derivative of wrt (i=0-n) to 0.

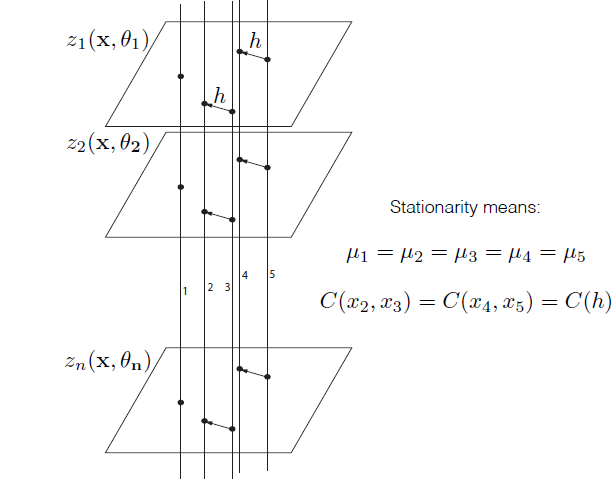
In matrix form all the n equations can be written as:

Value of will be replaced in the above **equation 1** to get the value of

# **Spatial relationships: Estimation and Modelling**

## Random Function Model

### Requirement of stationarity

A random function model is *loosely* stationary when the mean is constant and the covariance is translation invariant. (For full stationarity all the *central* moments has to be tested?)

In geostatistics, we mostly deal with Gaussian distribution (any other distribution is (normal score) transformed to Gaussian/Normal distribution). Gaussian distribution has only two central moments and when both are constant; we can say the Gaussian distribution is fully stationary.

Mathematically, the first order stationarity can be written as:

The second order stationarity can be defined as:

any function of two random variables distance apart is independent of the location and only dependent on the distance and direction (if anisotropic).

Covariance is the commonly used function relating the two variables.

## Spatial relationship

Spatial relationships are used to describe how neighbouring values are related. Three main relationships are: covariance, correlation coefficient & variogram. The dataset can either belong to one variable X spatially distributed or multiple variable X, Y spatially distributed.

### Covariance

In practice (on available sample data):

Where assumed to be 1st and 2nd order stationary

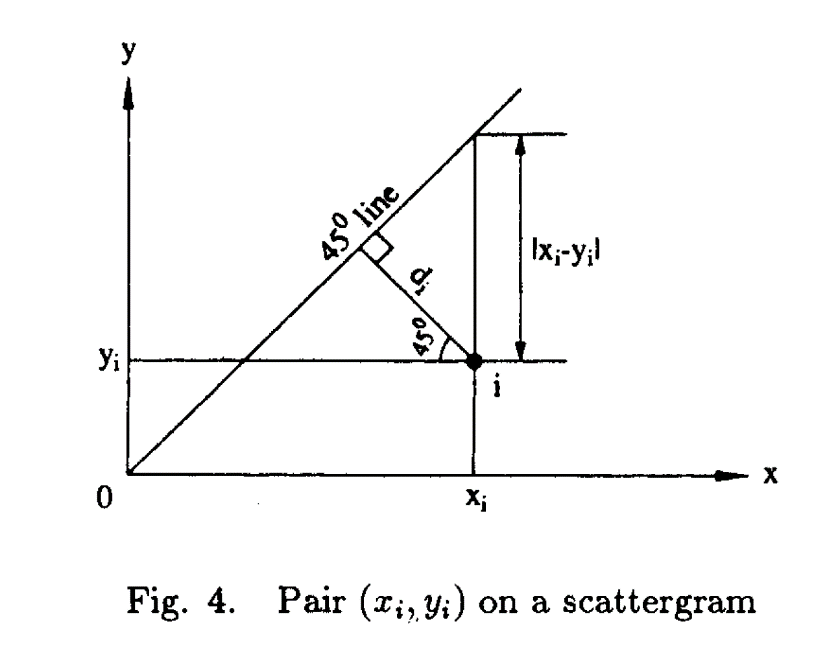
### Correlation Coefficient

For entire populations:

Where

### Variogram

#### What is variogram?

Semivariance (*moment of inertia*) is calculated, as shown in the scatterplot alongside:

1. Plot the variables in the scatterplot
2. Calculate the vertical distance from the x=y line
3. Now calculating the average of

Thus, the moment of inertia of the scattergram around e.g. the 45° line would be a characteristic of lack of dependence. This moment of inertia is called the "semi-variogram" of the set of data pairs and the variogram, is none other than the average of squared difference between the two components of each pair.

When the semivariance of each lag distance is plotted we get the variogram.

Note: Another way of looking at variogram. (From Kelkar, Perez)

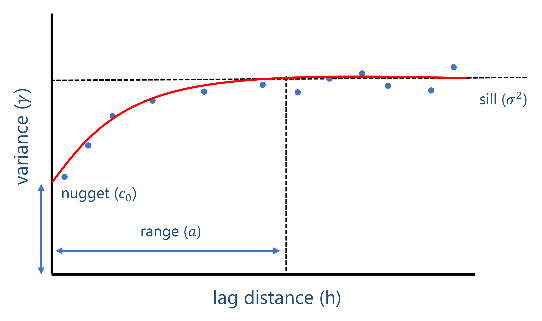
It is *half* of the *variance* of the difference of the value of the variable at two locations located distance apart.

#### Relationship between Variograms and Covariance

=

OR

Or variogram (**semivariance** between 2 points) = **sample variance**  – **covariance** between the 2 points. (at range, the semivariance reaches the max. sample variance)

Steps to get a variogram:

* 1. For each lag distance , the is calculated or just 1 point is plotted. (see diagram alongside)
  2. With changing lag distance lag distance , all the points are plotted to get the variogram.

#### Relationship between variogram and covariance (for non-ergodic systems)

For practical purposes (from the data set), the estimated variogram is:

**&**

**&**

Hence, by substituting the above variables in the equation above: (no stationarity principle applied)

The above equation can only turn into when the following stationarity holds true:



The above relations hold true only when we have sufficient pairs at each lag distance.

If we don’t have sufficient pairs, sample mean and variance can be different from lag mean and variance. In that case, we use the **non-ergodic variogram**.

Non-ergodic correlogram:

It accounts for the variation in the sample and lag variances.

## Estimation of Variogram

### Lack of Sufficient Pairs

#### Number of Pairs

For a fixed number of data points, the pairs keep decreasing with the increasing lag distance and so does the confidence in our data.

Thumb Rule: The variogram calculations should be restricted to *½ of the max distance* between any two data points within the ‘region of interest’ where the stationarity is assumed.

#### Lag Tolerance

Another alternative to ensure sufficient number of pairs for a given lag distance, is to describe a tolerance with respect to distance as well as direction/azimuth. (Because in a spatial distribution samples are not distributed at exactly)

Any point in the space can be taken into the calculation of the semivariance of 2 consecutive lag distances. E.g. a point come in the lag 32 and lag 42 calculations.

Thumb rule: Usually one should start variogram calculation with a small tolerance. If too many fluctuations, we will keep adjusting the tolerance until we get an interpretable structure. E.g., ratio of the range in the x-direction to y-direction is 5, which reduces to 2.4 once the directional tolerance is increased to 45 degrees.

### Instability of variogram

Since the estimated variogram is the arithmetic average of the squared differences of the two variables, any big differences when squared can have a significant impact on the semivariance value.

Two ways it can be reduced:

* Increasing the number of pairs at a given lag distance
* Remove certain extreme pairs (with very large differences). Cross plot the two variables (for a certain lag distance

### Influence of Outliers

#### Log Transform

Using the logarithmic values of the samples

#### Power Transform

Using the power values of the samples:

#### Rank Transform

All the values are set in ascending order and assigned the Rank, which are then used as values

#### Indicator Transform

The indicator transform allows us to transform the continuous variable into a discrete variable.

##### Indicator Variable

Indicator variable

So indicators can take either 0 or 1. By setting different thresholds, we can get multiple indicator values.

###### Expected Value

Cumulative distribution function

###### Variance

Since can only take values of 0 or 1, .

So,

###### Non-centred covariance

One of the unique properties of the indicator variable is its ability to define the connectivity between two points for a given threshold. This information cannot be obtained by conventional variograms.

Non-centered covariance

e.g.

= 1 \*

In calculating the product of indicator variables, we must consider FOUR possibilities, of these ONLY ONE CASE will result in a value of 1 and rest all 0.

In reservoir description, knowing the connectivity or knowing the barrier (when both the values are below the threshold) is very important.

E.g. for knowing connectivity, we can define the variable as

Hence,

By substituting (a) in (b),

**Indicator variables in facies:** Indicator variables are also used for describing categorical or discrete variables, e.g. geological facies. We can define the indicator variable as:

Similarly, we can define the joint probability distribution as below:

In addition, we can define joint connectivity between the two locations with two different facies criteria,

This can be easily extended to Multipoint histogram (2F.b) for categorical variables.

Advantages:

* + On appropriately defining the threshold values and estimating indicator variogram at each threshold, we can examine how sample values are connected at each thresholds.
  + Indicators allows qualitative information like geological facies to be translated into an indicator function.
    - E.g. facies 1: (1,0,0) facies 2(0,1,0) facies 3: (0,0,1)

#### Normal Score Transform

### Biased Sampling

#### General Relative Variogram

#### Pairwise Relative Variogram

#### Non-ergodic Variogram

#### Non-ergodic correlogram

## Modelling Variograms

The primary purpose in estimating the variogram is to use this information to estimate values at unsampled locations.

### Modelling requirements

Two requirements for modelling:

* Use minimum number of parameters & models to model the variogram.
* Condition of positive definiteness, which ensures a unique solution is obtained during the estimation process. Positive definiteness can be satisfied in different dimensions. A model which satisfies positive definiteness in a higher dimension will definitely satisfy in lower dimensions.

### Models with Sills

Models which reach a constant value (sample variance) after a certain lag distance called **range**.

#### kriging modelsNugget-Effect Model

#### Spherical Model

#### Exponential Model

#### Gaussian Model

#### Combination Model

### Models without sills

Models that doesn’t reach a constant value within the region of interest.

#### Fractional Gaussian Noise fGw Model

#### Fractional Brownian Motion fgw Model

### Hole Effect Models

Models showing a periodic behaviour (representing cyclical geological processes)

#### Cosine Model

### Anisotropic Model

#### Geometric Anisotropy

#### Zonal Anisotropy

## Cross-variograms

### Estimation of Cross variograms

**Cross variogram for 2 random variables X and Y** can be written as below:

And meaning cross-variogram is symmetric.

Cross-covariance:

For a lag of :

When the 1st order stationarity holds

Although 2nd terms of the & are equal, however first terms are not equal.

Hence, or cross-covariance is not symmetric. In some mining applications they are observed to be significantly different. However, in most reservoir characterization applications we assume that to be symmetric.

From :

OR

due to stationarity.

If cross covariance is assumed to be symmetric i.e. :

**In practice, cross-variogram is calculated as:** ()

Similar to conventional variogram,

and

and

Replace in :

In most instances we assume that,

Thus,

Replacing the definition of non-ergodic cross-variogram,

For and to be equal,

The local cross-covariance must be equal to the global cross-covariance i.e.

The local mean of first data point in a pair is equal to the local mean of the second data point of the pair, i.e. ,

**Markov-Bayes approximation for cross correlation.**

### Modelling of Cross-variograms

## Alternative methods of Spatial Relationships

### Modified Variograms

### Multipoint Histograms

# **Conventional Estimation Techniques**

## Preliminary Considerations

### Search Neighbourhood

#### Number of samples

A minimum of 4 is required to get an estimate and 12 samples is reasonable to get an estimate. Adding more points beyond the maximum 32 will not significantly improve the estimate.

#### Size of Neighbourhood

#### Shape of Neighbourhood

### Cross-validation

## Linear Kriging Procedures

The estimated value of the variable is linearly related to nearby samples.

### Simple Kriging

Value at the unsampled location can be estimated as: (similar to [**MVUE**](#_Inference_of_Parameters))

Real value: Estimated value:

Although the above relations are written as in random variable, but in actual they are the actual samples at the mentioned location.

#### Applying Unbiased condition

under the condition of 1st order stationarity

#### Applying Minimum Variance Condition

To minimize variance, we need to equate partial derivative of wrt .

In matrix form all the n equations can be written as:

Covariance values from the model variogram calculated earlier.

Value of will be replaced in the above **equation 1** to get the value of

#### Minimum variance relationship in terms of variogram/semivariance

From :

In matrix form all the n equations can be written as:

is required to calculate the values.

#### Calculating error variance

Multiplying with and summing over all the values of :

Replacing in to get as below:

has the maximum impact to calculate the error variance.

in terms of the variogram/semivariance values:

### Simple Block Kriging

Similar to point kriging, we start with the equation:

#### Unbiased condition:

#### Applying Minimum Variance Condition

Similar to simple kriging to minimize variance, we need to equate partial derivative of wrt .

#### Calculating error variance

Notice the similarity between point and block kriging expressions except for covariance values either within a block or between blocks.

### Ordinary Kriging

Similar to the Simple Kriging

Real value: Estimated value:

Although the above relations are written as in random variable, but in actual they are the actual values of the samples recorded at the mentioned location.

#### Applying Unbiased condition

Because we don’t know the value of the local mean, we can force to 0, which results in:

This is the *unbiased condition*.

#### Applying Minimum Variance Condition

We must minimize the variance by taking into account the constraint defined in . To achieve this we adopt the Lagrange multiplier method and define a function as below:

To minimize variance, we need to equate partial derivative of wrt and .

In matrix form all the n equations can be written as:

Covariance values and are calculated from the model variogram.

From , we are able to calculate and thus

#### Calculating error variance

Multiplying with and summing over all the values of :

Replacing in to get as below:

##### Field Example: Ordinary Kriging Cross-validation

##### Field Example: Ordinary Kriging to generate Gross Thickness Maps

##### Field Example: Ordinary kriging to estimate original oil in place

##### Field Example: Block Estimation

### Cokriging

When simple kriging is extended to 2 variables. We assume they are principal variable and the covariable. The estimation equation can be written as:

Estimated value:

#### Applying Unbiased Condition

For *simple cokriging system*, is the equation of unbiased state.

If the means and are not known and we force to 0, will result in an *ordinary cokriging system* as below.

& Thus,

But one of the drawbacks is some of the will result in negative weights, thus resulting in negative estimate at some locations which doesn’t have any physical meaning. Hence, to avoid the negative weighing, we’ve another option as below:

&

In the above relation, requires the knowledge of the means and

Thus the unbiased relations are as below:

Replacing in ,

Replacing in , :

Depending, on which constraint is chosen or , we either proceed with or

#### Calculating Minimum variance (Ordinary Kriging\_1)

Using constraint and thus the relation:

The error variance equation must be minimized alongwith the constraints from:

Taking derivatives of F wrt to and and equating to 0 we obtain:

The above set of equations can be written as in the matrix format below:

Where will give the values of and .

#### Calculating Error variance (Ordinary Kriging\_1)

#### Calculating Minimum variance (Ordinary Kriging\_2)

By using constraint and thus the relation:

The above set of equations can be written as in the matrix format below:

#### Calculating Error variance (Ordinary Kriging\_2)

=

For option 2, we must know the means of the variables X and Y.

Similar set of formulae can be generated for a **simple cokriging algorithm**.

#### Calculating Minimum variance (Simple Kriging)

Using the unbiased condition, where we can calculate when are calculated and are known already (condition for Simple Kriging).

Also the Estimated value:

#### Calculating Error variance (Simple Kriging)

#### Collocated cokriging

#### Generalized cokriging procedure

##### Unbiasedness

##### Minimum Variance

##### Field Example: Cokriging to estimate permeability using Initial Potential

### Universal Kriging

#### Unbiased condition

#### Minimum variance

#### Generalized Covariance Function

#### Kriging with external drift

##### Unbiasedness

##### Minimum variance

##### Field Example: Universal Kriging to Estimate Top of reservoir

##### Field Example: Universal Kriging with an external drift to estimate permeability using porosity trend

## Non-linear Kriging techniques

### Log-Normal Kriging

### Multi-Gaussian Kriging

### Indicator kriging

#### Indicator Kriging to generate Permeability Thickness Map

### Probability Kriging

## Estimation of uncertainty

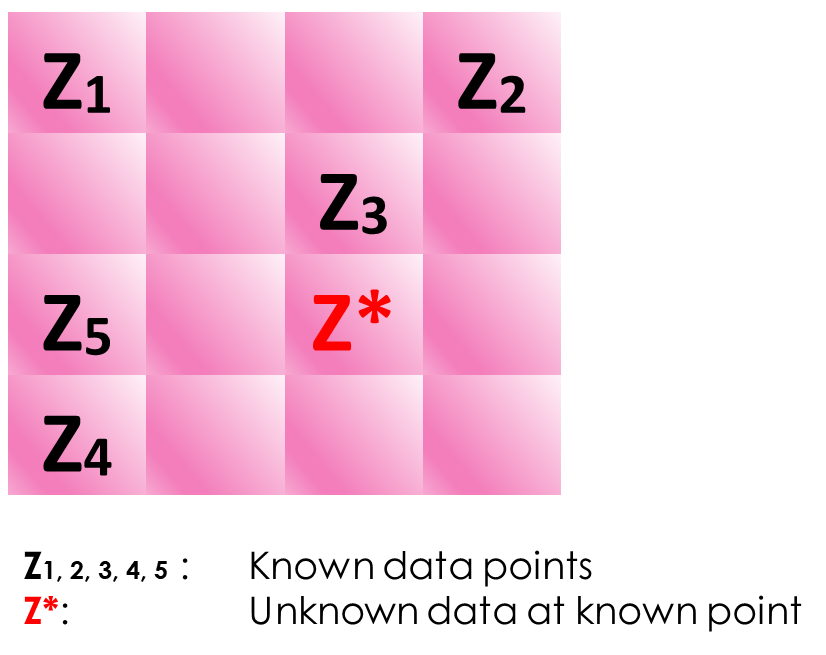
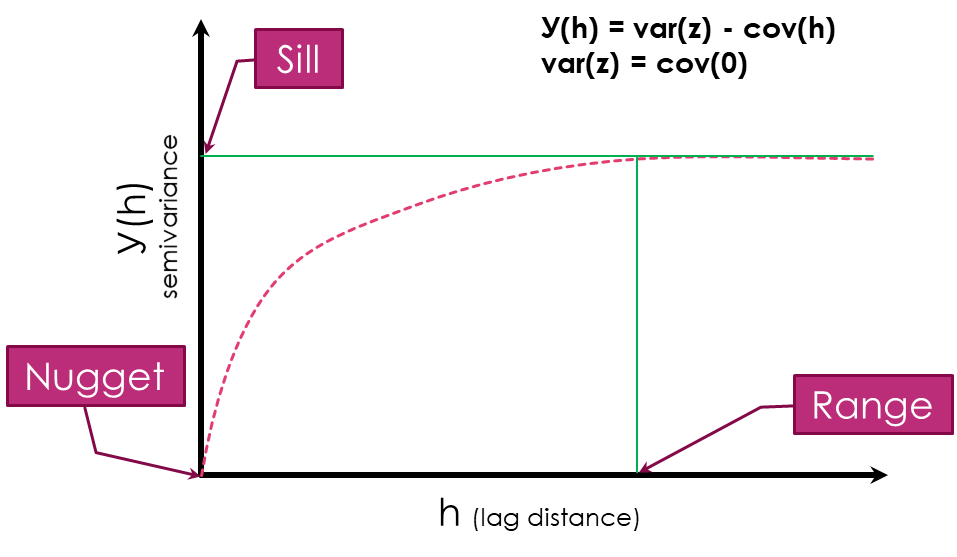
### Parametric Estimations

### Non-parametric Estimations

### Loss Functions

## Other Notes

### Simple Kriging (other books)

We have a few known data points Zi at Xi location and we need to calculate the Z\* value at X\*. (Zi, i=1…n, are all the data points within the search radius and which are taken into account to calculate Z\*.)

**Steps:**

1. Variogram: Calculate the 2D variogram for the points. (as above, gives the spatial relationship between points).

However, here we are going to assume it to be isotropic, so we can use a single variogram.

1. [](https://www.google.com/url?sa=i&rct=j&q=&esrc=s&source=imgres&cd=&cad=rja&uact=8&ved=&url=https://www.123rf.com/photo_41723343_stock-vector-hand-drawn-creative-idea-lightbulb-.html&psig=AOvVaw08A8rgeyVA0hdmrp8Bm32R&ust=1545725730885220)Search Neighbourhood: It specifies which data are to be included in the weighted moving average.

Thumb rule: Search radius ≥ Range from variogram (in the same direction). If using a uniform search radius in all directions, use the search radius comparable to the range of the long axis of the variogram.

1. Kriging Estimator:

= weights (dependent on the covariances between the data points)

= values at location

1. Estimation variance: The estimation variance has to be minimum. Substituting the value of (the kriging estimator) from last step, the estimation variance is a function of the λs.

For to be minimum, partially differentiating w.r.t. and equating to 0.

We get a set of equations, as below:

Where k= 1… N. When put together in a matrix as below.

1. Kriging Matrix:

**K** is calculated only once.

**k** is calculated for each unknown value (e.g. porosity) at a known location.

K. λ=k

Covariance between known data points & ONE unknown data point

Covariance between known data points

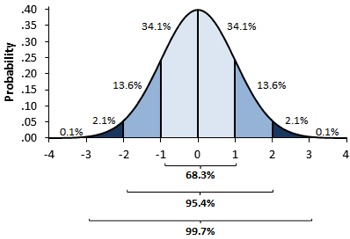
is calculated as below:

1. Measure and read the corresponding value from the variogram,
2. Using the formula , calculate the covariance matrices and solve for λ.
3. RESULT: Solved unknown values λ:
4. RESULT: Simple Kriging (SK) variance:

(Because)

At the end of the kriging, any kind of software will generate two maps:

1. Map of the Kriged values
2. Map of the ‘Simple Kriging’ variance.

**Salient points about Simple Kriging:**

1. It requires knowledge of the population mean, which may not be known in practice without prior assumptions. (Kelkar, Perez p.98)
2. At the time of the calculation, we need to specify a mean for the population, as against the OK.
3. If we do ‘normal score transform’ to the data, the variance/sill is always going to be 1.
4. Calculated Kriged weights always declusters the data, giving lower weightage to each data-point in a dense cluster and high weightage to each data-point in a sparse cluster.

### Ordinary Kriging (other books)

Ordinary Kriging (OK) assumes the data to be unbiased. Hence, we need to equate the ‘bias equation’ to 0.

Considering each of the points to be random variables . (instead of fixed values )

1. **Bias Equation:**

Error: ()

Bias = Mean of the error = ⇒ = =

⇒

Therefore, the above additional conditional needs to be satisfied for the Ordinary Kriging.

: Estimated value at the location

: True value at the same location

1. **Another version of the Error variance**: (Isaaks, Srivastava)
2. Error of the j-th estimate:

where

1. Average error for all **k** estimated points or Bias:

The *mean* of the error distribution is often referred to as **bias.**

When the distribution is assumed to be ‘unbiased’,

Assuming stationarity i.e.

1. Variance of error:

1. We switch from estimating the variable itself to estimating the random function from which reality is “a realization”. Thus, assuming the ‘random variable model’, we have a n random variables and calculating the (n+1)th variable.

V(x) or : Random function at location x.

: Estimated value

: Actual value

* 1. Our estimated value at the know location:
  2. Error or residual:
  3. Variance of the error/residual:

[**See Appendix**](https://clouddrive.aramco.com.sa/personal/dassx0a/Documents/OFFICE_DOCS/UPDC%20Courses/GeoStats%20Notes/Variogram%20&amp;%20Kriging.docx#_Appendix:)

Where

Combining the three terms,

We can solve the problem, by differentiating w.r.t. wi and solving them. However, earlier as per the unbiasedness condition, . Thus, we have (n+1) equations to solve n unknowns. Such constrained optimization problem can be solved by the ‘technique of Lagrange parameters’.

* 1. Technique of Lagrange parameters:

We introduce an extra unknown ‘Lagrange parameter µ’, without disturbing the equality

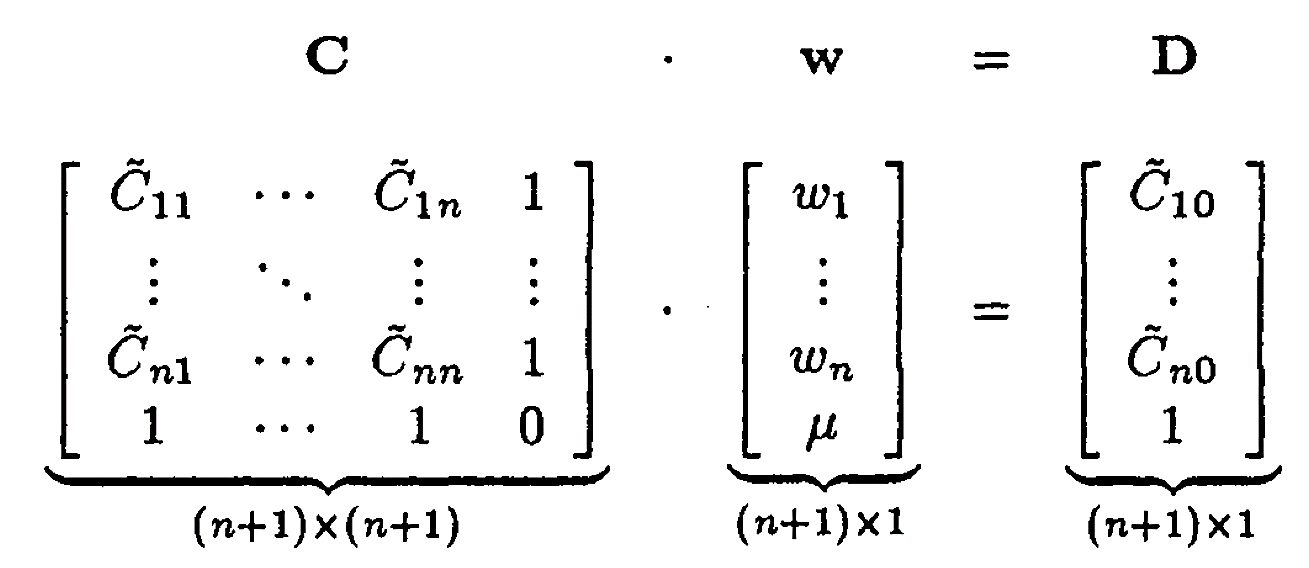
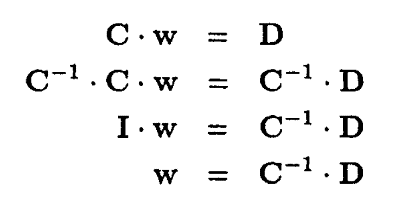
**0**

)

* 1. Now by partially differentiating up to first derivatives *w.r.t. µ* and equating to 0
* , reproduces the ‘unbiasedness condition’.
  1. Now by partially differentiating *w.r.t. wi* up to first derivatives

Adding up all the components,

* 1. Putting together all the (n+1) equations in a matrix, often referred to as ‘ordinary kriging matrix’

 ⇒

The above equations gives us the result of the (n+1) unknowns, wi & µ. Covariances are calculated from the variogram, using the formula:

RESULT: Solved unknown values wi:

\*\*Similarly, we calculate the for all the unknown points.

* 1. Calculate the minimized error variance or **Ordinary Kriging Variance**:

… from (d)

from (f)

Multiplying by ,

Substituting the above in (d)

**Other salient points about Ordinary Kriging:**

1. Ordinary Kriging (OK) is more robust than the Simple Kriging (SK), since it nullifies the local mean each time it calculates and work with the residuals.
2. The variance estimation error is independent of the sample values.

# Conditional Simulation Techniques

## Definition

## Distinguishing Features

### Sample Variability

### Sample Spatial Relationship

### Quantification of Uncertainty

## Simulation Methods

### Grid-Based Simulation Methods

### Object Based Simulation Techniques

# Grid Based Simulation Methods

## Sequential Conditional Simulation Methods

### Single Variable Simulation

#### Transformation into a New Domain

#### Random Path Selection

#### Estimation and Unsampled Location

#### Back Transform

#### Field Example: Sequential Gaussian Simulation

#### Field Example: Sequential Indicator Simulation for continuous variable

#### Field Example: Sequential Indicator Simulation for discrete variable

### Sequential Cosimulation

#### Transformation into a new domain

#### Estimation and modelling variograms

#### Selection of a random path

#### Simulation of Attributes

#### Sequential Gaussian Cosimulation

#### Bayes Rule Cosimulation

#### Proportion Curves Cosimulation

#### Field Example: Sequential Gaussian Cosimulation

## Probability Field Simulation

#### Probability Field Simulation

## Simulated Annealing

### Background

### Simulation Algorithm

#### Initial Distribution

#### Objective Function

#### Spatial relationship

#### Cumulative Distribution Function

#### Correlation between variables

#### Well Test Permeability

#### Production Data

#### Interchange Mechanism

## Simulation Process

### Input Parameters

#### Simulated Annealing Simulation

# Object Based Simulation Techniques

## Marked Point Process technique

### Simulation Process

### Properties of objects

### Conditioning Data

### Interactions among objects

## Criteria for choosing object modelling

### Depositional Environment

### Conditioning Data

### Proportion of objects

### Object Dimensions

### Seismic Data

### Available data

## Hybrid conditional simulation approaches

### Multistep approach using object modelling

### Multistep hybrid approach using object modelling and pixel based techniques