

Short Course on

Introduction to Statistics and Data Analysis in Geology 2022

Francesco Salvini

Francesco.salvini@uniroma3.it

When: February 21-25, 2022 9:00-13:30

Where: Aula L6 & Teams Platform

Register by e-mail to: Paola.Tuccimei@uniroma3.it

Our **GeoQuTe Lab** web site for
stuff downloading:

<http://host.uniroma3.it/progetti/fralab/>

This (very short) course on *Introduction to Geostatistics* is divided into 5 parts:

PART 1 - Introduction and Basic Statistics (Day 1)

PART 2 - Analysis of Sequence of Data (Day 2)

PART 3 - Analysis of Map Data (Day 3)

PART 4 - Analysis of Multivariate Data (Day 4-5)

**PART 5 – Something beyond... (Polymodal distributions,
Fractal distributions, Montecarlo approach...) (Day 5)**

The Statistics Program used in the class and a pdf version of the slides are available
at our **GeoQuTe Lab** web site:

<http://host.uniroma3.it/progetti/fralab/>

Abridged Program:

PART 1 - Introduction and Basic Statistics

Why Statistics

- Goals
- Assumptions
- Statistics versus Geostatistics
- The non-random sampling bias
- The non-ripetitive sampling
- The advent of the Computer Age

Type of Data Measurements

- Nominal Scale, Ordinal scale, Interval scale, Ratio scale
- Precision of Measurements
- One more bias...different measure precision
- Values and Errors
- Systematic error (biasing factors): Experimental, Instrumental, Operator, Data Variability Error
- Reliability
- Uniqueness of “Best Solution”
- Interpolation versus Extrapolation
- Reliability test... Attractor Analysis
- Evaluation of the Measure Error, GIGO (Garbage in, garbage out)

Population Statistics

- Probability
- Binomial distribution
- Normal Distribution
- Population and Sample Statistics

Parameters and Statistics

- Mean, Mode, Median
- Variance and Standard Deviation
- Covariance
- Correlation Coefficient r
- The Closure Problem

Testing Normal Population

- Standard Normal Form, The Critical Region
- Inference: Null Hypothesis / Alternative Hypothesis
- Error Alpha and error Beta
- Z test
- Student’s t test (comparing mean), Two way Student’s t test
- F test (comparing variance)
- Analysis of Variance (ANOVA), Two way ANOVA
- Chi Square Test
- Standardization

PART 2 - Analysis of Sequence of Data

Geological sequences
Equal spacing and interpolation
Runs Tests
Regression Analysis
Filtering sequences
Autocorrelation and Cross-correlation
Auto-association and Cross-association
Fourier Series
Event series
Transition Matrices, Markov Chain, Substitutability Analysis

PART 3 - Analysis of Map Data

Distributions of points
Contouring
Trend surfaces
4D Trend surfaces
Double Fourier Analysis (FFT)
Moving Averages and Kriging
Map Comparison

PART 4 – Analysis of Multivariate Data

Multiple Regression
Multivariate Statistics
Discriminant Functions
Cluster Analysis
PCA analysis
Factor Analysis, R-mode, Q-Mode Analyses

PART 5 – Something Beyond...

Multiple population data
Polymodal distribution analysis
Fractal Distribution
Montecarlo Approach to complex populations, Attractor Analysis