INFN School of statistics 2017

Ischia, 7-11 May 2017 — Hotel Continental Terme

Timetable	7/5/2017 Sunday	8/5/2017 Monday	9/5/2017 Tuesday	10/5/2017 Wednesday	11/5/2017 Thursday
9:00-9:30					
9:30-10:00		Lecture 2.1	Lecture 4.1	Lecture 5.1	Lecture 7.1
10:00-10:30					
10:30-11:00		Coffee break	Coffee break	Coffee break	Coffee break
11:00-11:30					
11:30-12:00		Lecture 2.2	Lecture 4.2	Lecture 5.2	Lecture 7.2
12:00-12:30	Registration				
12:30-13:00			Lunch		
13:00-13:30	Lunch	Lunch		Lunch	Departure
13:30-14:00					
14:00-14:30					
14:30-15:00					
15:00-15:30	Lecture 1.1	Lecture 3.1		Lecture 6.1	
15:30-16:00					
16:00-16:30	Coffee break	Coffee break	Excursion	Coffee break	
16:30-17:00					
17:00-17:30	Lecture 1.2	Lecture 3.2		Lecture 6.2	
17:30-18:00					
18:00-18:30					
18:30-19:00					
19:00-19:30					
19:30-20:00					
20:00-20:30	Dinner	Dinner	Dinner	Dinner	
20:30-21:00					
21:00-21:30					

Lecture programme

Lectures 1.1, 1.2 — Probability theory
Roger Barlow (Univ. of Huddersfield)
Introduction to probability theory
Random variables: discrete and continuous distribution, cumulative function
Most popular discrete and continuous statistical distributions
Multiple random variables: covariance and correlation coefficient
Distribution moments: expected value and variance
Conditional probability and Bayes theorem
Bayesian approach to probability
Bayesian inference

Lectures 2.1, 2.2 — Statistical methods
Glen Cowan (Royal Holloway, London)
Parameter estimates, properties of estimators
Maximum likelihood method
Error propagations
The error ellipse and its properties
Neyman belt and confidence intervals
Pearson and Neyman chi-squares
Combination of more measurements

Lectures 3.1, 3.2 — Multivariate analysis with complex networks
Antonio Scala (CNR, Rome)
Null models
Centralities and Rankings
Reconstructing Networks from Partial Data
Communities & Clustering
Standard datasets & Bipartite Networks
Community finding

Lectures 4.1, 4.2 — Hypothesis testing and interval estimation
Olaf Behnke (DESY, Hamburg)
Hypothesis testing
Goodness of fit
Frequentist and Bayesian interval estimation

Lectures 5.1, 5.2 — Multivariate analysis, 1
Ilya Narsky (MathWorks, USA)
Types of machine learning
Classification vs regression, numeric, nominal and ordinal variables
Decision boundaries and probability estimation
Learning curve, overtraining, various types of regularization
Cross-validation, stratified and non-stratified partitioning

Boosting as stagewise additive modeling: AdaBoost, GentleBoost and LogitBoost
Estimation of posterior class probabilities for binary classification by boosting
Extensions of boosting for multiclass problems and data with class imbalance
Bagging and random forest
Choosing decision tree parameters for boosting and random forest

Lectures 6.1, 6.2 — Multivariate analysis, 2
Andrey Ustyuzhanin (Yandex, Russia)
Artificial neutral networks
Deep learning
Examples of ML applications to the Physics

Lectures 7.1, 7.2 — Statistical software tools, with hands-on session
Mario Pelliccioni (INFN Torino)
Overview of the main statistical tools
Strong/weak point of the main tools
RooFit
RooStats
Usage examples code demonstrations