1. Record Nr. UNINA9910138994803321

Autore	Narsky Ilya
Titolo	Statistical analysis techniques in particle physics : fits, density estimation and supervised learning / / Ilya Narsky and Frank C. Porter
Pubbl/distr/stampa	Weinheim : , : Wiley-VCH, , [2014] ©2014
ISBN	3-527-67729-1 3-527-67732-1 3-527-67731-3
Descrizione fisica	1 online resource (461 p.)
Disciplina	530.4
Soggetti	Particles (Nuclear physics) - Statistical methods Physics Condensed matter
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Statistical Analysis Techniques in Particle Physics; Contents; Acknowledgements; Notation and Vocabulary; 1 Why We Wrote This Book and How You Should Read It; 2 Parametric Likelihood Fits; 2.1 Preliminaries; 2.1.1 Example: CP Violation via Mixing; 2.1.2 The Exponential Family; 2.1.3 Confidence Intervals; 2.1.4 Hypothesis Tests; 2.2 Parametric Likelihood Fits; 2.2.1 Nuisance Parameters; 2.2.2 Confidence Intervals from Pivotal Quantities; 2.2.3 Asymptotic Inference; 2.2.4 Profile Likelihood; 2.2.5 Conditional Likelihood; 2.3 Fits for Small Statistics 2.3.1 Sample Study of Coverage at Small Statistics2.3.2 When the pdf Goes Negative; 2.4 Results Near the Boundary of a Physical Region; 2.5 Likelihood Ratio Test for Presence of Signal; 2.6 sPlots; 2.7 Exercises; References; 3 Goodness of Fit; 3.1 Binned Goodness of Fit Tests; 3.2 Statistics Converging to Chi-Square; 3.3 Univariate Unbinned Goodness of Fit Tests; 3.3.1 Kolmogorov-Smirnov; 3.3.2 Anderson-Darling; 3.3.3 Watson; 3.3.4 Neyman Smooth; 3.4 Multivariate Tests; 3.4.1 Energy Tests; 3.4.2 Transformations to a Uniform Distribution; 3.4.3 Local Density Tests; 3.4.4 Kernel-based Tests

3.4.5 Mixed Sample Tests3.4.6 Using a Classifier; 3.5 Exercises; References; 4 Resampling Techniques; 4.1 Permutation Sampling; 4.2 Bootstrap; 4.2.1 Bootstrap Confidence Intervals; 4.2.2 Smoothed Bootstrap; 4.2.3 Parametric Bootstrap; 4.3 Jackknife; 4.4 BCa Confidence Intervals; 4.5 Cross-Validation; 4.6 Resampling Weighted Observations; 4.7 Exercises; References; 5 Density Estimation; 5.1 Empirical Density Estimate; 5.2 Histograms; 5.3 Kernel Estimation; 5.3.1 Multivariate Kernel Estimation; 5.4 Ideogram; 5.5 Parametric vs. Nonparametric Density Estimation; 5.6 Optimization
5.6.1 Choosing Histogram Binning5.7 Estimating Errors; 5.8 The Curse of Dimensionality; 5.9 Adaptive Kernel Estimation; 5.10 Naive Bayes Classification; 5.11 Multivariate Kernel Estimation; 5.12 Estimation Using Orthogonal Series; 5.13 Using Monte Carlo Models; 5.14 Unfolding; 5.14.1 Unfolding: Regularization; 5.15 Exercises; References; 6 Basic Concepts and Definitions of Machine Learning; 6.1 Supervised, Unsupervised, and Semi-Supervised; 6.2 Tall and Wide Data; 6.3 Batch and Online Learning; 6.4 Parallel Learning; 6.5 Classification and Regression; References; 7 Data Preprocessing
7.1 Categorical Variables7.2 Missing Values; 7.2.1 Likelihood Optimization; 7.2.2 Deletion; 7.2.3 Augmentation; 7.2.4 Imputation; 7.2.5 Other Methods; 7.3 Outliers; 7.4 Exercises; References; 8 Linear Transformations and Dimensionality Reduction; 8.1 Centering, Scaling, Reflection and Rotation; 8.2 Rotation and Dimensionality Reduction; 8.3 Principal Component Analysis (PCA); 8.3.1 Theory; 8.3.2 Numerical Implementation; 8.3.3 Weighted Data; 8.3.4 How Many Principal Components Are Enough?; 8.3.5 Example: Apply PCA and Choose the Optimal Number of Components
8.4 Independent Component Analysis (ICA)

Sommario/riassunto

Modern analysis of HEP data needs advanced statistical tools to separate signal from background. This is the first book which focuses on machine learning techniques. It will be of interest to almost every high energy physicist, and, due to its coverage, suitable for students.