

Kurzbeschreibung der Blockkurse

Nonparametric and Resampling Methods (Lukas Meier, ETH Zürich und Diego Kuonen, Statoo Consulting)

Nonparametric tests, randomization tests, jackknife and bootstrap, asymptotic approximations of estimators.

Robust Regression (Andreas Ruckstuhl, ZHAW Winterthur)

Influence function, breakdown point, regression M-estimation, regression MM-estimation, robust inference, covariance estimation with high breakdown point, application in principal component analysis and linear discriminant analysis.

Nonlinear Regression (Andreas Ruckstuhl, ZHAW Winterthur)

Nonlinear regression model, estimation methods, approximate tests and confidence intervals, profile t -plot, profile traces, parameter transformation, prediction and calibration.

Statistics for Survival Data (Alain Hauser, Google)

The primary purpose of a survival analysis is to model and analyze time-to-event data; that is, data that have as a principal endpoint the length of time for an event to occur. This block course introduces the field of survival analysis without getting too embroiled in the theoretical technicalities. We will present some frequently used parametric models and methods, including accelerated failure time models; and the nonparametric procedures which include the Kaplan-Meier estimate of survival and the Cox proportional hazards regression model. The statistical tools treated are applicable to data from medical clinical trials, public health, epidemiology, engineering, economics, psychology, and demography as well.

Nichtparametrische Regression (Martin Mächler, ETH Zürich)

Fokus ist die nichtparametrische Schätzung von Wahrscheinlichkeitsdichten und Regressionsfunktionen. Diese neueren Methoden verzichten auf einschränkende Modellannahmen wie "lineare Funktion". Sie benötigen eine Gewichtsfunktion und einen Glättungsparameter. Schwerpunkt ist eine Dimension, mehrere Dimensionen und Stichproben von Kurven werden kurz behandelt.

Spatial Statistics (Madlene Nussbaum, Utrecht)

Many phenomena are spatial, hence geographical analysis of data is relevant to a wide range of domains like geology, soil science, agronomy, ecology, forestry, hydrology, climatology, human geography or epidemiology. When analyzing such data one might be interested in exploring their structure (i.e., dependence on explanatory variables, autocorrelation), in distribution of point patterns or in producing spatial prediction to obtain continuous maps. The course provides an introduction to geostatistical methods like kriging approaches and machine learning with a focus on spatial prediction and model evaluation.

Repeated Measures (Meta-Lina Spohn, Zurich Data Scientists)

Analyse von Verlaufskurven und anderen wiederholten Messungen. Analyse von Profil-Eigenschaften, Compound-Symmetry und intra-class Korrelation, Split-Plot Modell, Kovarianzstruktur von Messungen mit zeitlicher oder räumlicher Distanz, Wachstumskurven, (ev.) multivariate Varianzanalyse, mehrere Behandlungen an derselben Versuchseinheit (Cross-Over Versuche).

High-Dimensional Statistics (Nicolas Städler, Roche)

Kurs über die Analyse hochdimensionaler Daten mit Fokus auf Vorhersageprobleme und das Feature Assessment. Teil 1: Lineare Regression im hochdimensionalen Kontext; Überanpassung, Vorhersage und das Bias-Varianz Dilemma; Modellselektion; Ridge und Lasso Regularisierung. Teil 2: Logistische Regression und Regularisierung; Klassifikation mittels Entscheidungsbäumen, Random Forest und AdaBoost; Multiples Testen, p-Wert-Anpassung, Varianz Schrumpfung.

Applied Bayesian Statistics (Sylvain Robert, Google)

We will learn how to describe business/scientific problems as probabilistic models, apply Bayes rules to draw inference from data, and use the probabilistic programming language STAN to obtain samples from posterior distributions. On the way we will fit linear models both for continuous and categorical outcomes, and explore techniques to deal with hierarchical structures in the data. There will be examples of applications from various fields: insurance, meteorology, marketing, etc.

Stichproben-Erhebungen (Tobias Schoch, FHNW)

Die Elemente einer Stichproben-Erhebung werden erklärt. Die wichtigsten klassischen Stichprobenpläne (einfach und geschichtete Zufallsstichprobe) mit ihren Schätzern sowie Schätzverfahren mit Hilfsinformationen und der Horvitz-Thompson Schätzer werden eingeführt. Datenaufbereitung, Antwortausfälle und deren Behandlung, Varianzschätzungen sowie Analysen von Stichprobendaten werden diskutiert.

Statistical Analysis of Financial Data (Andreas Ruckstuhl und Marcel Dettling, ZHAW Winterthur)

Univariate Finanzzeitreihen: Preise, Returns, Einfaches Modell: Unabhängigkeit, Schätzung und Limiten im Einsatz, Spezielle Zeitreihen-Modelle (GARCH, ...), Prognose. Univariate Risikomasse: Sigma, Value at Risk / Expected Shortfall, Implikationen aus den Modellen auf diese Grössen (bzw. deren Schätzungen). Multivariate Analyse bzw. Portfoliooptimierung: Rendite und Risiko, Schwierigkeiten in praktischer (multivariater) Umsetzung, Copulas, Praktische Umsetzung und Limiten, Aktuelle Situation und Alternativen.

Deep Learning (Beate Sick, ZHAW Winterthur und Oliver Dürr, Hochschule Konstanz)

Dieser Kurs führt in probabilistisches deep learning (DL) ein. DL ist die Methode der Wahl für komplexe Daten, wie Bilder usw. Hier wird deep learning als Fortführung bekannter probabilistischer Modelle, wie logistische Regression im Framework von Maximum Likelihood und Bayes Methoden behandelt. Sie lernen verschiedene Architekturen neuronaler Netze (NN) kennen (z.B. fully connected und convolutional neural networks) und erfahren, wie Sie die passende NN-Architektur für Ihre Aufgabe auswählen. Sie werden lernen, verschiedene Verteilungen der Zielgrösse (wie Gauss, Poisson oder Multinomial) für die jeweilige Aufgabe zu modellieren. Sie werden praktische Erfahrungen im Aufbau probabilistischer DL-Modelle sammeln, lernen, wie man sie abstimmt, und wie man das Trainingsverfahren kontrolliert.