Institut für
Mathematische Stochastik

Georg–August–Universität

Lecture series

Introduction to Nonparametric Goodness-of-Fit Testing

by

Assistant Professor Dr. Farida Enikeeva, University of Poitiers

October 21, 22 and 24, 25, 2019; 12:15 pm-13:45 pm

Venue: Institute for Mathematical Stochastics, Goldschmidtstr. 7, SR 5.101

Abstract

The problem of hypotheses testing aims at construction of a decision rule that allows to distinguish between two given hypotheses. In parametric problems the hypotheses are completely described by a finite-dimensional parameter. In case of a one-dimensional parameter the classical result of Neyman and Pearson provides an optimal test for the problem of testing two simple hypotheses. This test, that is often called likelihood ratio test, minimizes the type II error in the class of all tests of a given significance level. In a classical course on parametric statistics the case of testing composite hypotheses is then treated by using the maximum likelihood ratio test that leads, in some situations, to (asymptotically) uniformly most powerful tests.

In the multidimensional parameter case it is difficult to define a uniformly most powerful test. In this case we can define a *minimax* optimal test that minimizes the (maximal over the set of alternatives) type II error over the class of all tests of a given significance level. Another approach to multidimensional testing is to impose some prior distributions on the sets of parameters corresponding to the null and the alternative hypotheses, and to find the best test of a given level that minimizes the average type II error with respect to these priors. These two approaches as well as the main results related to the minimax testing will be considered in Lecture 1.

The classical problem of non-parametric goodness-of-fit testing deals with the hypotheses that the underlying function describing the data (density, distribution function, the signal in the white noise model) belongs to some class of functions. In Lectures 2 and 3 we will consider the problems of goodness-of-fit testing on the signal observed in the Gaussian white noise model. We will construct asymptotically minimax tests in the problem of testing the simple null hypotheses of absence of the signal ($H_0 : f = 0$) against quadratic or Sobolev-type alternatives. We will study the performance of the χ^2 (chi-squared) tests, χ^q tests with $q \in (0,+\infty)$ and the suprema tests. Several approaches to the construction of lower bounds of minimax testing risk will be considered.

Lecture 4 will be devoted to some problems of high-dimensional testing with sparse alternatives. We will consider the problem of detection of exactly $p \asymp n^{1-\beta}$ non-zero components in the mean of a high-dimensional Gaussian vector of dimension $n \to \infty$. We will show that the rate of testing depends on the sparsity coefficient β of the problem. We will construct two different tests corresponding to two sparsity zones and find their minimax detection rate. Some related problems as the detection of a sparse submatrix of a high-dimensional noisy matrix and the detection of the change in mean in a high-dimensional Gaussian vector will be considered.

The research visit of Dr. Enikeeva is partly funded by the Emmy-Noether-Gastprogramm.