TEST : AN OFFICIAL JOURNAL OF THE SPANISH SOCIETY OF STATISTICS AND OPERATIONS RESEARCH, ISSN 1133-0686
Volume 31, number 2 (june 2022)

Abstract
In the literature, most works of the specification tests focus on the problem with one-dimensional response or fixed multidimensional responses. In this paper, we develop a new specification test for the parametric models with non-stationary regressor under multidimensional setup, where the dimension of responses may tend to infinity, which fills a gap in the literature. The theoretical results about the asymptotic properties of the proposed test are studied and the optimal rate of the local departure under the alternative hypothesis is also given which ensures the models underpinning by the null and alternative hypotheses can be differentiated. Some simulation studies are done to evaluate the performance of the proposed test with the finite sample. Besides, a real data example based on the US aggregate consumers' consumption data is employed to illustrate the performance. The results of simulation studies and real data analysis both demonstrate the efficiency of our proposed method.

Abstract
M-type smoothing splines are a broad class of spline estimators that include the popular least-squares smoothing spline but also spline estimators that are less susceptible to outlying observations and model misspecification. However, available asymptotic theory only covers smoothing spline estimators based on smooth objective functions and consequently leaves out frequently used resistant estimators such as quantile and Huber-type smoothing splines. We provide a general treatment in this paper and, assuming only the convexity of the objective function, show that the least-squares (super-)convergence rates can be extended to M-type estimators whose asymptotic properties have not been hitherto described. We further show that auxiliary scale estimates may be handled under significantly weaker assumptions than those found in the literature and we establish optimal rates of convergence for the derivatives, which have not been obtained outside the least-squares framework. A simulation study and a real-data example illustrate the competitive performance of non-smooth M-type splines in relation to the least-squares spline on regular data and their superior performance on data that contain anomalies.

Abstract
We propose a method of testing a shift between mean vectors of two multivariate Gaussian random variables in a high-dimensional setting incorporating the possible dependency and allowing $p>n$. This method is a combination of two well-known tests: the Hotelling test and the Simes test. The tests are integrated by sampling several dimensions at each iteration, testing each using the Hotelling test, and combining their results using the Simes test. We prove that

this procedure is valid asymptotically. This procedure can be extended to handle non-equal covariance matrices by plugging in the appropriate extension of the Hotelling test. Using a simulation study, we show that the proposed test is advantageous over state-of-the-art tests in many scenarios and robust to violation of the Gaussian assumption.

Spatial distribution of invasive species: an extent of occurrence approach P. 416-441

Alberto Rodríguez-Casal, Paula Saavedra-Nieves

Abstract

Ecological Risk Assessment faces the challenge of determining the impact of invasive species on biodiversity conservation. Although many statistical methods have emerged in recent years in order to model the evolution of the spatio-temporal distribution of invasive species, the notion of extent of occurrence, formally defined by the International Union for the Conservation of Nature, has not been properly handled. In this work, a novel and flexible reconstruction of the extent of occurrence from occurrence data will be established from nonparametric support estimation theory. Mathematically, given a random sample of points from some unknown distribution, we establish a new data-driven method for estimating its probability support S in general dimension. Under the mild geometric assumption that S is r–convex, the smallest r–convex set which contains the sample points is the natural estimator. A stochastic algorithm is proposed for determining an optimal estimate of r from the data under regularity conditions on the density function. The performance of this estimator is studied by reconstructing the extent of occurrence of an assemblage of invasive plant species in the Azores archipelago.

Single-index composite quantile regression for ultra-high-dimensional data P. 443-460

Rong Jiang, Mengxian Sun

Abstract

Composite quantile regression (CQR) is a robust and efficient estimation method. This paper studies CQR method for single-index models with ultra-high-dimensional data. We propose a penalized CQR estimator for single-index models and combine the debiasing technique with the CQR method to construct an estimator that is asymptotically normal, which enables the construction of valid confidence intervals and hypothesis testing. Both simulations and data analysis are conducted to illustrate the finite sample performance of the proposed methods.

New characterization-based exponentiality tests for randomly censored data P. 461-487

Marija Cuparić, Bojana Milošević

Abstract

Recently, the characterization-based approach for the construction of goodness of fit tests has become popular. Most of the proposed tests have been designed for complete i.i.d. samples. Here, we present the adaptation of the recently proposed exponentiality tests based on equidistribution-type characterizations for the case of randomly censored data. Their asymptotic properties are provided. Besides, we present the results of wide empirical power study including the powers of several recent competitors. This study can be used as a benchmark for future tests proposed for this kind of data.

Tests for circular symmetry of complex-valued random vectors P. 488-518

Norbert Henze, Pierre Lafaye De Micheaux, Simos G. Meintanis

Abstract

We propose tests for the null hypothesis that the law of a complex-valued random vector is circularly symmetric. The test criteria are formulated as $L2$-type criteria based on empirical characteristic functions, and they are convenient from the computational point of view. Asymptotic as well as Monte Carlo results are presented. Applications on real data are also reported. An R package called CircSymTest is available from the authors.

Abstract
We propose a new test to validate the assumption of homoscedasticity in a functional linear model. We consider a minimum distance measure of heteroscedasticity in functional data, which is zero in the case where the variance is constant and positive otherwise. We derive an explicit form of the measure, propose an estimator for the quantity, and show that an appropriately standardized version of the estimator is asymptotically normally distributed under both the null (homoscedasticity) and alternative hypotheses. We extend this result for residuals from functional linear models and develop a bootstrap diagnostic test for the presence of heteroscedasticity under the postulated model. Moreover, our approach also allows testing for "relevant" deviations from the homoscedastic variance structure and constructing confidence intervals for the proposed measure. We investigate the performance of our method using extensive numerical simulations and a data example.

Abstract
Given a finite set of independent random variables, assume one can observe their sum, and denote with s its value. Efron in 1965, and Lehmann in 1966, described conditions on the involved variables such that each of them stochastically increases in the value s, i.e., such that the expected value of any non-decreasing function of the variable increases as s increases. In this paper, we investigate conditions such that this stochastic monotonicity property is satisfied when the assumption of independence is removed. Comparisons in the stronger likelihood ratio order are considered as well.