Annals of the Institute of Statistical Mathematics, ISSN 0020-3157
Volume 67, number 3 (june 2015)

Abstract
In this paper, we study integrated regression techniques to check the adequacy of a given model in the context of selection-biased observations. We introduce integrated regression in this setting, providing not only a suitable statistic for enabling a model checking test, but also a bootstrap distributional approximation to carry out the test. We also address the behaviour of the test under different alternatives showing that this behaviour is asymptotically the same for both selection-biased and non selection-biased data. The technique is illustrated with a simulation study and a data analysis based on a real situation that shows the performance of the method and how selection bias affect both estimation and inference.

Abstract
In this article, we propose a new test for additivity in nonparametric quantile regression with a high-dimensional predictor. Asymptotic normality of the corresponding test statistic (after appropriate standardization) is established under the null hypothesis, local and fixed alternatives. We also propose a bootstrap procedure which can be used to improve the approximation of the nominal level for moderate sample sizes. The methodology is also illustrated by means of a small simulation study, and a data example is analyzed.

Abstract
We study nonparametric robust tail coefficient estimation when the variable of interest, assumed to be of Weibull type, is observed simultaneously with a random covariate. In particular, we introduce a robust estimator for the tail coefficient, using the idea of the density power divergence, based on the relative excesses above a high threshold. The main asymptotic properties of our estimator are established under very general assumptions. The finite sample performance of the proposed procedure is evaluated by a small simulation experiment.

Abstract
We determine the joint limiting distribution of adjacent spacings around a central, intermediate, or an extreme order statistic TeX of a random sample of size TeX from a continuous distribution TeX . For central and intermediate cases, normalized spacings in the left and right neighborhoods are asymptotically i.i.d. exponential random variables. The

associated independent Poisson arrival processes are independent of TeX . For an extreme TeX , the asymptotic independence property of spacings fails for TeX in the domain of attraction of Fréchet and Weibull (TeX) distributions. This work also provides additional insight into the limiting distribution for the number of observations around TeX for all three cases.

On the equivariance criterion in statistical prediction P. 541-555

Haojin Zhou - Tapan K. Nayak

Abstract

This paper presents a general development of the basic logic of equivariance for a parametric point prediction problem. We propose a framework that allows the set of possible predictions as well as the losses to depend on the data and then explore the nature and properties of relevant transformation groups for applying the functional and formal equivariance principles. We define loss invariance and predictive equivariance appropriately and discuss their ramifications. We describe a structure of equivariant predictors in terms of maximal invariants and present a method for deriving minimum risk equivariant predictors. We explore the connections between equivariance and risk unbiasedness and show that uniquely best risk unbiased predictors are almost equivariant. We apply our theoretical results to some illustrative examples.

Probabilistic properties of second order branching process P. 557-572

Akanksha S. Kashikar - S. R. Deshmukh

Abstract

The classical BGW process assumes first order dependence, whereas many real life datasets exhibit a second or higher order dependence. Further, in some situations, there is a need for a model which allows for simultaneous reproduction by a parent and its offspring. This paper proposes a second order branching process model to accommodate such situations and discusses its probabilistic properties such as extinction probability and limiting behaviour of the generation sizes. Estimation of offspring means and growth rate are also discussed. This model is further used to model the swine flu data for Pune, India, and La-Gloria, Mexico.

The sinh-arcsinhed logistic family of distributions: properties and inference P. 573-594

Arthur Pewsey - Toshihiro Abe

Abstract

The sinh-arcsinh transform is used to obtain a flexible four-parameter model that provides a natural framework with which to perform inference robust to wide-ranging departures from the logistic distribution. Its basic properties are established and its distribution and quantile functions, and properties related to them, shown to be highly tractable. Two important subfamilies are also explored. Maximum likelihood estimation is discussed, and reparametrisations designed to reduce the asymptotic correlations between the maximum likelihood estimates provided. A likelihood-ratio test for logisticness, which outperforms standard empirical distribution function based tests, follows naturally. The application of the proposed model and inferential methods is illustrated in an analysis of carbon fibre strength data. Multivariate extensions of the model are explored.

Smooth change point estimation in regression models with random design P. 595-619

Maik Döring - Uwe Jensen

Abstract

We consider the problem of estimating the location of a change point $\theta 0$ in a regression model. Most change point models studied so far were based on regression functions with a jump. However, we focus on regression functions, which are continuous at $\theta 0$. The degree of smoothness $\varphi 0$ has to be estimated as well. We investigate the

consistency with increasing sample size n of the least squares estimates $(\hat{\theta}n,\hat{\varphi}n)$ of $(\theta 0,\varphi 0)$. It turns out that the rates of convergence of $\hat{\theta}n$ depend on $\varphi 0$: for $\varphi 0$ greater than $1/2$ we have a rate of $n\sqrt{}$ and the asymptotic normality property; for $\varphi 0$ less than $1/2$ the rate is $n1/(2\varphi 0+1)$ and the change point estimator converges to a maximizer of a Gaussian process; for $\varphi 0$ equal to $1/2$ the rate is $n\cdot \ln(n)\text{------}\sqrt{}$. Interestingly, in the last case the limiting distribution is also normal