Séminaire Mathématique de Béjaïa (LaMOS)
Volume 11, 2012, pp. 77 - 78

18

On the Estimation of the Mode by Orthogonal Series

Nora SAADI et Smail ADJABI

LAboratoire de Modélisation et d'Optimisation des Systèmes LAMOS
Université de Béjaïa 06000, Algérie.

Résumé The simplicity and good performance of orthogonal series density estimator have found favor with several authors, and have been applied to several different areas (harmonic analysis, signal processing, image compression, physical analysis, functional statistics ...). This paper addresses the problem of estimating the mode of a density function under non parametric conditions based on an orthogonal series. We give a rigorous, theoretical account of the estimator's properties(bias, variance, mean square error, convergence of the bias, convergence of the variance, convergence of the mean square error and convergence in probability). Our results show that some of the theoretical properties of Orthogonal series's estimator are similar to those of Parzen's estimator. For example, under the condition that the density has two bounded derivatives in a neighborhood of the mode, and the density is uniformly continuous.

Mots clés : Mode estimation, Orthogonal series,Density estimation, Convergence in probability.

18.1 Introduction

The most fundamental problem in statistics is the determination of the unknown density of probability function $h(.)$ and the mode x_0. Several works have been devoted to the estimation of the mode under non parametric conditions. Most of these works require regularity conditions such as the existence of the density function, the density being Lipschitz or being twice differentiable, for example. Given a sample set $X_1, ..., X_n$ of independent identically distributed variables, each with the unknown probability density $h(.)$ and mode x_0. We want to estimate x_0 from this sample set. The classic nonparametric estimator of the mode x_0 is the kernel estimator introduced by Parzen (1962). The last quinquennial has seen an increased interest and numerous papers are dealing with the problem of the mode. To quote only a few of them, Mokkadem (2006) studied mode rate and large deviation supper bounds for the kernel mode estimator. Hermann and Ziegler (2004) obtained rates of nonparametric estimation of the mode in absence of smoothness assumptions. However, the most of the available studies concentrate on kernel estimator introduced by Parzen (1962), and the majority of studies (see for example Asselin de beauville (1978), Gasser, Hall and Presnell (1998)) have consisted of Monte Carlo experiments, with very little attention being paid to theoretical results. There does not even exist a description of theoretical properties of the mode estimators using orthogonal series. But, in many areas of mathematical analysis, the smoothness of a function is more readily determined by the behaviour of its Fourier series. our thesis in this paper is that the latter approach is natural and convenient when analyzing properties orthogonal series density and mode estimators. orthogonal series density estimators were introduced by Cencove (1962) and have since been applied to several different areas (harmonic analysis, signal processing, image compression, functional statistics ...). Orthogonal series estimators have found favor in applications because of the ease with which they can bean be calculated and for their good performance with the multivariate data (see Frey (1976)). In the last decade however, several papers have appeared which deal with the estimation of orthogonalnal expansion of densities . Some notable examples include (Bosq and Blank(2007), Herrick, Nason and Silverman (2001), and Efromovich (2010)). This method assumes that a square-integrable $h(.)$ can be represented as a convergent orthogonal series expansion, $h(x) = \sum_{k=0}^{\infty} a_k e_k(x)$, $x \in \Omega$, where $\{e_k(.)\}_{k=0}^{\infty}$ is a complete orthonormal system of functions on a set Ω of the real line that is, satisfying $\int_{\Omega} e_k(x)e_j(x)dx = \delta_{ij}$, $0 \leq k,\ j < \infty$, where δ_{ij} is the delta Kronecker and $\{a_k\}$ are Fourier coefficients defined by $a_k = \int_{\Omega} e_k(x)h(x)dx = \mathbb{E}[e_k(X)], k = 0, \ldots$ Orthonormal systems proposed for $\{e_k(.)\}_{k=0}^{\infty}$ are those with compact support (such as the Fourier, trigonometric, and Haar systems on $[0, 1]$, and Legendre systems on $[-1, 1]$) and those with unbounded support such as the Hermite system on $(-\infty, \infty)$, and Laguerre system on $[0, \infty[$, (see Walter (1994)). A discussion of trigonometric systems can be found in (Efromovich (2010)). Classical orthogonal polynomials, including Legendre, Gegenbauer, Jacobi, and Chebyshev, are a popular choice as well ; (see Walter (1994), Budzkis and Radavicius (2005)). Our aim in the present paper is to provide a rigorous account of several theoretical aspects of Orthogonal series mode estimator and we generalize the terme of the variance given by Asselin de beauville (1978) in trigonometric cases. In particular we prove the convergence in probability of the estimator. and we give its statistical and asymptotic properties(bias, variance, mean square error, convergence of the bias, convergence of the variance and convergence of the mean square error).

References

Asselin de beauville, J. (1978). Estimation non paramétrique de la densité de probabilité et du mode exemple de la distribution gamma. *Revue de Statistique Appliquée, (Tome26, N°3) :47-70.*

Cencov, N. (1962). Evaluation of an unknown distribution density from observations.*Soviet Math. Dokl, 3, 1559–1562*.

Bosq,D., Blanke, D. (2007). Inference and Prediction in Large Dimensions. *Wiley.*

Efromovich, S. (2010). Orthogonal series density estimation. *Interdisciplinary review : computational statistics, 2, 467-476.*

Fryer, M. (1976). A Review of Some Non-Parametric Estimators of Density Functions. *Journal of the Institute of Mathematics and Its Applications, 18, 371-380*.

Gasser,T., Hall, P. and Presnell, B. (1998). Nonparametric estimation of the mode of a distribution of random curves. *Journal of the Royal Statistical Society. 60, 681-691*.

Herrmann, E. and Ziegler, K. (2004). Rates of consistency for nonparametric estimation of the mode in absence of smoothness assumptions. *Statistics and Probability Letters, 68 :359-368.*

Herrick, D., Nason, G., Silverman, B. (2001). Some new methods for wavelet density estimation. *Sankhya, A63, 94-411*.

Parzen, E. (1962). On estimation of a probability density function and mode. Ann. Math. Stat., 33(3) :1065-1076.

Mokkadem, A., Pelletier, M., Worms, J. (2006). Alarge deviation supper bound for the kernel mode estimator. *Theory Probab.Appl.50, 153-165.*

Rudzkis R, Radavicius M.(2005). Adaptive estimation of distribution density in the basis of algebraic polynomials. *TTheory Probab Appl , 49 :93-109*.

Walter, G. (1994). Wavelets and other Orthogonal Systems with Application. *London : CRC Press.*