Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 1829 – 1837
journal of statistical planning and inference
www.elsevier.com/locate/jspi

The LIL for the Bickel–Rosenblatt test statistic

Han-Ying Liang[a,*], Bing-Yi Jing[b]

[a]Department of Applied Mathematics, Tongji University, Shanghai 200092, PR China
[b]Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

Received 29 April 2005; accepted 6 June 2006
Available online 14 August 2006

Abstract

Let $\{X_n, n \geqslant 1\}$ be independent identically distributed random variables with common density function, f. We are interested in testing the hypothesis $\mathrm{H}_0 : f = f_0$ vs. $\mathrm{H}_1 : f \neq f_0$, where f_0 is a given density. In this note, we discuss the law of the iterated logarithm (LIL) of the Bickel–Rosenblatt test statistic under fixed alternatives. These results are of particular importance if the null hypothesis cannot be rejected. Also the proof of the results shows that the test statistics are of different order under the null hypothesis and the alternative hypothesis.

MSC: Primary 62G07; secondary 62M07

Keywords: Kernel density estimator; Bickel–Rosenblatt test statistic; Law of the iterated logarithm; Fixed alternative

1. Introduction

Let $\{X_n, n \geqslant 1\}$ be independent identically distributed random variables with common density function, f, with respect to the Lebesgue measure on $\mathbf{R}$. The classical density estimator of f based on sample $(X_1, X_2, \ldots, X_n)$ is defined as

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right), \quad x \in \mathbf{R},$$

where h_n is a sequence of real positive numbers tending to zero as n tends to infinity and K is a kernel function on $\mathbf{R}$. The asymptotic normality of the estimator, f_n, has been studied by many authors. For instance, related to the L_p-norms of the density estimator, $p \geqslant 1$, we refer to Hall (1984), Csörgö and Horváth (1988), Giné et al. (2003), and Csörgö et al. (1991) for the censorship case and Horváth (1991) for the multivariate case.

In this note, we are interested in tests of fit pertaining to f. We consider the following hypothesis:

$$\mathrm{H}_0 : f = f_0 \quad \text{versus} \quad \mathrm{H}_1 : f \neq f_0, \tag{1}$$

* Corresponding author.
E-mail addresses: hyliang83@yahoo.com (H.-Y. Liang), majing@ust.hk (B.-Y. Jing).

0378-3758/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2006.06.036

where f_0 is a given density. Many tests have been proposed for this problem (see Lehmann, 1986; d'Agostino and Stephens, 1986). Bickel and Rosenblatt (1973) proposed a test based on the L_2 distance between a nonparametric density estimate and a parametric fit of the density:

$$T_n = \int_{\mathbf{R}} (f_n - K_h * f_0)^2(x)\,\mathrm{d}x \tag{2}$$

or

$$\tilde{T}_n = \int_{\mathbf{R}} (f_n - f_0)^2(x)\,\mathrm{d}x, \tag{3}$$

where $K_h(\cdot) = h^{-1}H(\cdot/h)$, $h := h_n$ and $f_1 * f_2$ is the convolution of the functions, f_1 and f_2. Then, the null hypothesis will be rejected for large values of T_n or $\tilde{T}_n$.

Under H_0, the asymptotic normality of T_n and $\tilde{T}_n$ has been studied extensively. See, for example, Bickel and Rosenblatt (1973), Lee and Na (2002), Bachmann and Dette (2004), Chebana (2004, 2006) and Tenreiro (2006). Actually, there have been attempts to extend the Bickel–Rosenblatt test to dependent processes, such as those by Takahata and Yoshihara (1987) and Neumann and Paparoditis (2000), who have considered the case when the X_i's are dependent random variables under suitable mixing conditions. In those papers, the main results are related to the limiting distribution of the test statistics. However, there are few results that consider the exact convergence order for the test statistics.

There was no result regarding the law of the iterated logarithm (LIL) for the integrated squared deviation of a kernel density estimator until the work by Giné and Mason (2004). Motivated by the work of Giné and Mason (2004), we provide a more refined analysis of the Bickel–Rosenblatt test through a discussion of the LIL of the test statistics under fixed alternatives of the form

$$\int_{\mathbf{R}} (f - f_0)^2(x)\,\mathrm{d}x > 0. \tag{4}$$

These results are of particular importance if the null hypothesis cannot be rejected (cf. Berger and Delampady, 1987; Sellke et al., 2001). Also the proof of main results shows that the test statistics are of a different order under the null and the alternative hypothesis.

This paper is organized as follows. In Section 2, we give the main results and make some remarks. The proofs of the main results are presented in Section 3.

Throughout the paper, let $\|\cdot\|_v$ denote the total variation norm and $\|\cdot\|_2$ the L_2 norm with respect to the Lebesgue measure on $\mathbf{R}$; $A_n \asymp B_n$ means that $0 < \liminf_n A_n/B_n \leqslant \limsup_n A_n/B_n < \infty$.

2. The main results

To state our results, we first introduce some assumptions.

(A.1) The density function, f, satisfies $\int_{\mathbf{R}} f^p(x)\,\mathrm{d}x < \infty$ for some $p > 2$.

(A.2) The kernel, K, is a measurable function such that

$$\|K\|_v < \infty,\ \|K\|_r < \infty \quad \text{for } r = 1, 2,\ \int_{\mathbf{R}} \sup_{|y| \geqslant |x|} |K(y)|\,\mathrm{d}x < \infty \quad \text{and} \quad \int_{\mathbf{R}} K(x)\,\mathrm{d}x = 1.$$

(A.3) The bandwidth, $h_n > 0$, satisfies

$$h_n \searrow 0,\ h_n \asymp n^{-\delta} \quad \text{for some } \delta \in (0, 1/3)$$

and there exists an increasing sequence of positive constants, $\lambda_k, k \geqslant 1$, satisfying $\lambda_{k+1}/\lambda_k \to 1$ and $\log\log \lambda_k/\log k \to 1$, as $k \to \infty$, such that

$$h_n \text{ is constant for } n \in [\lambda_k, \lambda_{k+1}), \quad k \in \mathbf{N}.$$

(A.4) The kernel, K, satisfies

$$K \geqslant 0, \quad \int_{\mathbf{R}} xK(x)\,\mathrm{d}x = 0, \quad \int_{\mathbf{R}} x^2K(x)\,\mathrm{d}x := 2k < \infty.$$