ELSEVIER

Statistics & Probability Letters 40 (1998) 215–226

STATISTICS & PROBABILITY LETTERS

A test for singularity[1]

Attila Frigyesi *, Ola Hössjer

Institute for Matematisk Statistik, Lunds University, Box 118, 22100 Lund, Sweden

Received March 1997; received in revised form February 1998

Abstract

The L^p-norm and other functionals of the kernel estimate of density functions (as functions of the bandwidth) are studied as means to test singularity. Given an absolutely continuous distribution function, the L^p-norm, and the other studied functionals, will converge under some mild side-constraints as the bandwidth decreases. In the singular case, the functionals will diverge. These properties may be used to test whether data comes from an arbitrarily large class of absolutely continuous distributions or not.

AMS classification: primary 60E05; secondary 28A99; 62G07

Keywords: Singularity; Absolute continuity; Kernel density estimation

1. Introduction

In this paper we aim to test whether a given set of data $(=\{X_1,\dots,X_n\})$, with a marginal distribution μ, stems from an absolutely continuous distribution function or from a singular one. A reference to this problem is by Donoho (1988).

Consider a probability measure μ on $(\mathbb{R}^p, \mathscr{B}(\mathbb{R}^p))$, where $\mathscr{B}(\mathbb{R}^p)$ is the Borel σ-algebra on $\mathbb{R}^p$. The measure μ is said to be singular with respect to the Lebesgue measure λ if there exists a measurable set A such that $\mu(A^c)=0$ and $\lambda(A)=0$. If we only require $\mu(A)>0$ and $\lambda(A)=0$, then μ is said to have a singular part. If μ does not have a singular part, then it is absolutely continuous with respect to λ.

When μ is absolutely continuous the density $f=\mathrm{d}\mu/\mathrm{d}\lambda$ is a well-defined function. An estimate of f is

$$\hat{f}_n(x)=\frac{1}{nh_n^p}\sum_{i=1}^{n}K\left(\frac{X_i-x}{h_n}\right),$$

* Corresponding author. E-mail: atilla.frigyesi@matstat.lu.se.

[1] Research sponsored in part by Office of Naval Research Grant N00014 93 1 0043.

0167-7152/98/$ – see front matter
PII S0167-7152(98)00087-X

the kernel density estimator. The bandwidth h_n is a sequence of positive numbers, and K is a Borel measurable function (kernel), satisfying $\int K(t)\,\mathrm{d}t = 1$. A good reference on kernel density estimation is Wand and Jones (1995).

We will also make use of $\hat{f}_n$, when μ has a singular part. In this case $\hat{f}_n$ is an absolutely continuous estimate of a generalized function.

Suppose $H:[0,\infty)\to[0,\infty)$ is a function such that

$$\tilde{H}(x) = H(x)/x$$

is continuous and non-decreasing on the positive real line, tending to infinity as x does. We will study the following family of functionals:

$$J(\hat{f}_n) = \tilde{H}^{-1}\left(\int H(\hat{f}_n(x))\,\mathrm{d}x\right),$$

where $\tilde{H}^{-1}$ is the inverse of $\tilde{H}$. (In case $\tilde{H}$ is not strictly increasing, we let $\tilde{H}^{-1}$ denote, say, the right continuous inverse.) If Y is a random variable with density function g, then

$$J(g) = \tilde{H}^{-1}(E(\tilde{H}(g(Y)))),$$

so $J(g)$ measures the size of $g(Y)$. Possible choices of $H(x)$ are x^2 (or more generally $x^{1+\varepsilon}$, $\varepsilon>0$) and $x(\log x)_+$.

In Section 3, we prove (under some mild side constraints) that as $n\to\infty$,

$$J(\hat{f}_n) \xrightarrow{p} J(f) < \infty,$$

when the data that $\hat{f}_n$ is based on comes from an absolutely continuous distribution with density f such that $J(f)$ is finite. In Section 2, we prove (under some other mild side constraints) that as $n\to\infty$,

$$J(\hat{f}_n) \xrightarrow{\text{a.s.}} \infty,$$

when the data that $\hat{f}_n$ is based on comes from a distribution with a singular part. (see Fig. 1).

Using these results we may devise a "test for singularity" by investigating the behavior of $J(\hat{f}_n)$ (or equivalently $\int H(\hat{f}_n)\,\mathrm{d}x$) as $n\to\infty$ and $h_n\to 0$. If $J(\hat{f}_n)$ diverges we may say that the data that $\hat{f}_n$ is based on comes from either a singular distribution or an absolutely continuous distribution with $J(f)=\infty$. Given any $\mu \ll \lambda$, we may find an H such that $J(f)<\infty$, so the class of absolutely continuous distributions for which $J(\hat{f}_n)$ diverges can be made arbitrarily small.

The basic idea is to let H magnify the "spikes" of $\hat{f}_n$, that occur when the distribution is singular and the number of observations grows.

2. Singular distribution functions: the case $J(\hat{f}_n) \xrightarrow{\text{a.s.}} \infty$

We need the following lemma for the main result in this section, and B^δ will denote

$$\{x;\ \inf_{y\in B} |x-y| \leqslant \delta\}.$$

Lemma 2.1. *Assume $\mu(A)>0$ for some $A\subseteq \mathbb{R}^p$ with $\lambda(A)=0$. Then there exists $B\subseteq A$ such that $\mu(B)>0$ and $\lambda(B^\delta)\to 0$ as $\delta\to 0$.*

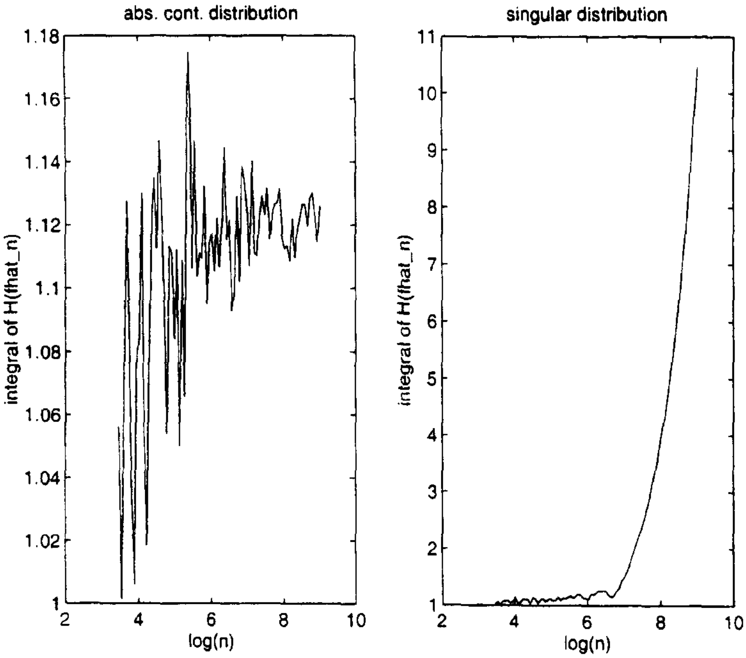

Fig. 1. The behavior of $J(\hat{f}_n)$ for different values of n. In the absolutely continuous case (uniform distribution of data $on[0,1]$) $J(\hat{f}_n)$ converges, while in the singular case (uniform distribution on the rational numbers $1/100,\dots,99/100$) $J(\hat{f}_n)$ diverges. $H(\hat{f}_n)=\hat{f}_n^2$ and $h_n=8/n$. Note that we resample for every value of n.

Proof. It is sufficient to find a set B with the following property:

(†) To every $\varepsilon>0$ there is a set $G_\varepsilon \supseteq B$, such that $\lambda(G_\varepsilon)<\varepsilon$ and G_ε is a finite union of (closed) boxes.

Provided $\varepsilon\leqslant 1$, all boxes in G_ε must have sides $\leqslant 1$. Let N_ε denote the number of boxes in G_ε. We then have

$$\lambda(B^\delta)\leqslant\lambda(G_\varepsilon^\delta)\leqslant\varepsilon+N_\varepsilon[(1+2\delta)^p-1]\leqslant 2\varepsilon,$$

provided that δ is small enough. As ε was chosen arbitrarily, the lemma will follow from (†).

To prove (†), we may, without loss of generality, assume $\mu(A)=1$. Construct $A\supseteq A_1\supseteq A_2\supseteq\cdots$ inductively according to the following: $A=A_0$, $A_k=A_{k-1}\cap\Omega_k$, $k\geqslant 1$, where Ω_k is a finite union of (closed) boxes with $\lambda(\Omega_k)\leqslant 2^{-k}$ and $\mu(A_{k-1}\backslash\Omega_k)\leqslant 2^{-(k+1)}$. Put $B=\bigcap_{k=1}^\infty A_k$. Then $\mu(A_{k-1}\backslash A_k)\leqslant 2^{-(k+1)}\Rightarrow \mu(A_k)\geqslant 1-\sum_1^k 2^{-(i+1)}>\frac{1}{2}\Rightarrow\mu(B)\geqslant\frac{1}{2}$. Since $B\subset\Omega_k$, $\forall k$, we can choose $G_\varepsilon=\Omega_k$ if $2^{-k}<\varepsilon$.

It now remains to construct Ω_k given A_{k-1}. We can find an open set $U_k\supset A_{k-1}$ with $\lambda(U_k)<2^{-k}$ since $\lambda(A_k)=0$. According to Vitali's covering theorem (see Cohn, 1980) there exists a sequence of (closed) disjoint boxes such that $\mu(A_{k-1}\backslash\bigcup_{n=1}^\infty C_{nk})=0$ and $C_{nk}\subset U_k\,\forall n$. Now choose N_k so that $\mu(A_{k-1}\backslash\bigcup_{n=1}^{N_k} C_{nk})\leqslant 2^{-(k+1)}$ and put $\Omega_k=\bigcup_{n=1}^{N_k} C_{nk}$. □

Let
(2.i) $\{X_i\}_{i\in\mathbb{Z}}$ be a stationary ergodic process with a marginal distribution μ,
(2.ii) $H\geqslant 0$,
(2.iii) $\tilde{H}(x)=H(x)/x$ be nondecreasing and tending to infinity as $x\to\infty$,
(2.iv) $K\geqslant 0$,

(2.v) $\int K\,\mathrm{d}x = 1$.

Theorem 2.1. *Under* (2.i)–(2.v), *if* μ *has a singular part, then*

$$\int H(\hat{f}_n)\,\mathrm{d}x \overset{\text{a.s.}}{\to} \infty$$

as $n \to \infty$ *and* $h_n \to 0$, *or equivalently* (*because of* (2.iii))

$$J(\hat{f}_n) \overset{\text{a.s.}}{\to} \infty.$$

Remark. We have not excluded the possibility of $J(\hat{f}_n) = \infty$, for some n. The requirement $\int K\,\mathrm{d}x = 1$ can be omitted. It is included here as it guarantees $\int \hat{f}_n\,\mathrm{d}x = 1$.

Proof. Assume $\operatorname{supp}(K) \subseteq [-1,1]^p$. According to our lemma there exists $B \subseteq \mathbb{R}^p$ such that $\mu(B) > 0$ and $\lambda(B^\delta) \to 0$ as $\delta \to 0$. Let $N = |\{X_i; X_i \in B\}|$. Put

$$\tilde{f}_n(x) = \frac{1}{nh_n^p} \sum_{X_i \in B} K\left(\frac{X_i - x}{h_n}\right).$$

We then have $\hat{f}_n \geqslant \tilde{f}_n \geqslant 0$, $\operatorname{supp}(\tilde{f}_n) \subseteq B^{h_n}$ and $\int \tilde{f}_n\,\mathrm{d}x = N/n \overset{\text{a.s.}}{\to} \mu(B)$. Let $D_n = \{x \in B^{h_n}; \tilde{f}_n(x) \leqslant N/(2n\lambda(B^{h_n}))\}$ and $D_n' = B^{h_n} \setminus D_n$. We then have

$$\int_{D_n} \tilde{f}_n\,\mathrm{d}x \leqslant \frac{N/n}{2\lambda(B^{h_n})}\lambda(D_n) \leqslant \frac{N/n}{2}$$
$$\Rightarrow \int_{D_n'} \tilde{f}_n\,\mathrm{d}x \geqslant \frac{N/n}{2}.$$

If $x \in D_n'$ then $\hat{f}_n(x) \geqslant \tilde{f}_n(x) > (N/n)/2\lambda(B^{h_n})$. Now,

$$\begin{aligned}\int H(\hat{f}_n)\,\mathrm{d}x &\geqslant \int_{D_n'} H(\hat{f}_n)\,\mathrm{d}x \\ &\geqslant \underline{H}\left(\frac{N/n}{2\lambda(B^{h_n})}\right)\int_{D_n'} \hat{f}_n\,\mathrm{d}x \\ &\geqslant \underline{H}\left(\frac{N/n}{2\lambda(B^{h_n})}\right)\frac{N/n}{2} \overset{\text{a.s.}}{\to} \infty,\end{aligned}$$

where $\underline{H}(x) = \inf_{y \geqslant x} H(y)/y$ for $x > 0$, and $\underline{H}(x) \to \infty$ as $x \to \infty$.

If $\operatorname{supp}(K) \nsubseteq [-1,1]^p$, then choose $C > 0$ such that $\int_{[-C,C]^p} K\,\mathrm{d}x > 0$. Put $\tilde{K} = K\chi_{[-C,C]^p}$ and

$$\tilde{f}_n(x) = \frac{1}{nh_n^p} \sum_{X_i \in B} \tilde{K}\left(\frac{X_i - x}{h_n}\right).$$

The rest of the proof is similar to the case $\operatorname{supp}(K) \subseteq [-1,1]^p$. $\square$

3. Absolutely continuous distribution functions: the case $J(\hat{f}_n) \overset{p}{\to} J(f) < \infty$

We begin with the main theorem of this section, followed by some lemmas that are necessary to the proof.

We will need the concept of ρ-mixing. A stationary process $\{X_i\}_{i\in\mathbb{Z}}$ is *ρ-mixing* if the ρ-mixing coefficients

$$\rho_n = \sup_{\substack{X\in\mathscr{F}_{-\infty}^0,\ Y\in\mathscr{F}_n^\infty \\ f,g}} \operatorname{Corr}[(f(X), g(Y)] \to 0,$$

as $n\to\infty$, for $n\in\mathbb{N}$. Here f,g range over all measurable functions with $Ef(X)^2<\infty$ and $Eg(X)^2<\infty$, and $\mathscr{F}_a^b$ denotes the σ-field induced by $\{X_i\}_{i=a}^b$ (see Hall and Heyde, 1980).

We will also be using the following assumptions:

(3.i) $\{X_k\}_{k\in\mathbb{Z}}$ is a stationary process with an absolutely continuous marginal distribution μ on $(\mathbb{R}^p, \mathscr{B}(\mathbb{R}^p))$ with compact support and a density f,

(3.ii) $\operatorname{supp}(K)$ is compact, $K\geqslant 0$, $\int K^2\,\mathrm{d}t<\infty$, and $\int K\,\mathrm{d}t=1$,

(3.iii) $h_n\to 0$, $nh_n^p\to\infty$, and $\sum_{k=0}^{n-1}\rho_k/(nh_n^p)\to 0$, when $n\to\infty$,

(3.iv) $H\geqslant 0$, $H(0)=0$,

(3.v) H is convex,

(3.vi) $H(x)/x^2\leqslant C_0$, $\forall x\in[1,\infty)$,

(3.vii) $\exists C_1>0$ s.t. $H(2x)\leqslant C_1 H(x)$, $\forall x\geqslant 0$,

(3.viii) $J(f)=\int_{\mathbb{R}^p} H(f(x))\,\mathrm{d}x<\infty$.

Theorem 3.1. *Under* (3.i)–(3.viii):

$$\int H(\hat{f}_n)\,\mathrm{d}x \xrightarrow{L_1} \int H(f)\,\mathrm{d}x,$$

and if $\tilde{H}$ is strictly increasing at $\tilde{H}^{-1}(\int H(f)\,\mathrm{d}x)$,

$$J(\hat{f}_n) \xrightarrow{P} J(f).$$

Proof. Write $\bar{J}(f)=\int H(f)\,\mathrm{d}x$. We may assume $\operatorname{supp}(\mu)\subseteq[\varepsilon,1-\varepsilon]^p$, for some $\varepsilon>0$ so that $\operatorname{supp}(\hat{f}_n)\subseteq[0,1]^p$ for n large enough. Let $\bar{f}_n(x)=\mathrm{E}[\hat{f}_n(x)]$. Now,

$$\bar{J}(\hat{f}_n)-\bar{J}(f)=[\bar{J}(\hat{f}_n)-\bar{J}(\bar{f}_n)]+[\bar{J}(\bar{f}_n)-\bar{J}(f)],$$

so that

$$E|\bar{J}(\hat{f}_n)-\bar{J}(f)|\leqslant E|\bar{J}(\hat{f}_n)-\bar{J}(\bar{f}_n)|+|\bar{J}(\bar{f}_n)-\bar{J}(f)|=T_1+T_2.$$

Lemma 3.1–3.3 give that

$$\begin{aligned} T_1 &\leqslant C_2\sqrt{\bar{J}(\bar{f}_n)}\mathrm{E}\|\hat{f}_n-\bar{f}_n\|_2 + C_3\mathrm{E}\|\hat{f}_n-\bar{f}_n\|_2^2 + C_4\mathrm{E}\|\hat{f}_n-\bar{f}_n\|_2 \\ &\leqslant C_2\sqrt{\bar{J}(f)}\mathrm{E}\|\hat{f}_n-\bar{f}_n\|_2 + C_3\mathrm{E}\|\hat{f}_n-\bar{f}_n\|_2^2 + C_4\mathrm{E}\|\hat{f}_n-\bar{f}_n\|_2 \to 0, \end{aligned}$$

as $n\to\infty$. Given a $L>0$ let

$$f^L(x)=f(x)\wedge L$$

and

$$\bar{f}_n^L(x)=\frac{1}{h_n^p}\int K\left(\frac{x-y}{h_n}\right) f^L(y)\,\mathrm{d}y.$$

We then get

$$T_2 \leqslant |\bar{J}(\bar{f}_n) - \bar{J}(\bar{f}_n^L)| + |\bar{J}(\bar{f}_n^L) - \bar{J}(f^L)| + |\bar{J}(f^L) - \bar{J}(f)| = T_{21} + T_{22} + T_{23}.$$

From Lemma 3.2 it follows that

$$T_{22} \leqslant C_2\sqrt{\bar{J}(f^L)}\|\bar{f}_n^L - f^L\|_2 + C_3\|\bar{f}_n^L - f^L\|_2^2 + C_4\|\bar{f}_n^L - f^L\|_2 \to 0,$$

as $n \to \infty$, as a consequence of L^2-continuity ($\bar{f}_n^L, f^L \in L^2(\mathbb{R}^p)$). Let $A_L = \{x;\ H(f(x)) > L\}$, giving that

$$T_{23} = \int [H(f(x)) - H(f^L(x))]\,\mathrm{d}x \leqslant \int_{A_L} H(f(x))\,\mathrm{d}x \to 0,$$

as $L \to \infty$, using Lebesgue's Dominated Convergence Theorem. For a fixed $x \in [0,1]^p$, put

$$\bar{K}(x,n) = \frac{1}{h_n^p}\int_{A_L} K\left(\frac{x-y}{h_n}\right)\mathrm{d}y$$

and

$$\eta^L(x) = f(x) - f^L(x).$$

We then get

$$\begin{aligned}
H(\bar{f}_n(x)) - H(\bar{f}_n^L(x)) &= H\left(\bar{f}_n^L(x) + \frac{1}{h_n^p}\int_{A_L} K\left(\frac{x-y}{h_n}\right)\eta^L\,\mathrm{d}y\right) - H(\bar{f}_n^L(x)) \\
&\leqslant \bar{K}\left[H\left(\bar{f}_n^L(x) + \frac{h_n^{-p}\int_{A_L} K\left(\frac{x-y}{h_n}\right)\eta^L\,\mathrm{d}y}{\bar{K}}\right) - H(\bar{f}_n^L(x))\right] \\
&\leqslant \bar{K}\left[H\left(L + \frac{h_n^{-p}\int_{A_L} K\left(\frac{x-y}{h_n}\right)\eta^L\,\mathrm{d}y}{\bar{K}}\right) - H(L)\right] \\
&\leqslant \frac{1}{h_n^p}\int_{A_L} K\left(\frac{x-y}{h_n}\right)H(L+\eta^L(y))\,\mathrm{d}y - \frac{1}{h_n^p}\int_{A_L} K\left(\frac{x-y}{h_n}\right)H(L)\,\mathrm{d}y,
\end{aligned}$$

using $0 \leqslant \bar{K} \leqslant 1$ and the convexity of H. Integration with respect to x gives

$$\begin{aligned}
T_{21} &= \int (H(\bar{f}_n(x)) - H(\bar{f}_n^L(x)))\,\mathrm{d}x \\
&\leqslant \frac{1}{h_n^p}\int_{A_L} K\left(\frac{x-y}{h_n}\right)[H(f^L(y) + \eta^L(y)) - H(f^L(y))]\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{A_L} [H(f^L(y) + \eta^L(y)) - H(f^L(y))]\,\mathrm{d}y = T_{23}.
\end{aligned}$$

That $T_2 \to 0$ when $n \to \infty$ follows by letting $n \to \infty$ and then $L \to \infty$.

Now assume $\tilde{H}$ is strictly increasing at $\tilde{H}^{-1}(\bar{J}(f))$. Then $\tilde{H}^{-1}$ is continuous at $\bar{J}(f)$, so the second half of the theorem easily follows from what we have just proved combined with Markov's inequality. □

Lemma 3.1. *$E\|\hat{f}_n - \bar{f}_n\|_2^2 \to 0$ as $n \to \infty$.*

Proof. For n large enough $\mathrm{supp}(\hat{f}_n) \subseteq [0,1]^p$, and $\mathrm{supp}(\bar{f}_n) \subseteq [0,1]^p$. Notice first that

$$E\|\hat{f}_n - \bar{f}_n\|^2 = \int_{[0,1]^p} \mathrm{Var}(\hat{f}_n)\,\mathrm{d}x,$$

but

$$\begin{aligned}\mathrm{Var}[\hat{f}_n(x)] &= \frac{n}{(nh_n)^{2p}}\mathrm{Var}\left[K\left(\frac{X_1 - x}{h_n}\right)\right] + \frac{2}{(nh_n)^{2p}}\sum_{i=1}^{n-1}(n - |i|)\mathrm{Cov}\left[K\left(\frac{X_1 - x}{h_n}\right), K\left(\frac{X_{1+i} - x}{h_n}\right)\right] \\ &\leqslant \frac{1}{(nh_n)^{2p}}\mathrm{Var}\left[K\left(\frac{X_1 - x}{h_n}\right)\right]\left(n + 2\sum_{i=1}^{n-1}(n - i)\rho_i\right) \\ &= \frac{1}{nh_n^{2p}}\mathrm{Var}\left[K\left(\frac{X_1 - x}{h_n}\right)\right]\left(1 + 2\sum_{i=1}^{n-1}\left(1 - \frac{i}{n}\right)\rho_i\right) \\ &\leqslant \frac{1}{nh_n^{2p}}\mathrm{Var}\left[K\left(\frac{X_1 - x}{h_n}\right)\right]\left(1 + 2\sum_{i=1}^{n-1}\rho_i\right),\end{aligned}$$

and

$$\mathrm{Var}\left[K\left(\frac{X_1 - x}{h_n}\right)\right] \leqslant E\left[K\left(\frac{X_1 - x}{h_n}\right)^2\right] = h_n^p \int K(t)^2 f(x + th_n)\,\mathrm{d}t.$$

Using Fubini's theorem we get

$$\begin{aligned}\int \mathrm{Var}[\hat{f}_n(x)]\,\mathrm{d}x &\leqslant \frac{1 + 2\sum_{i=1}^{n-1}\rho_i}{nh_n^p}\int K(t)^2 \int f(x + th_n)\,\mathrm{d}x\,\mathrm{d}t \\ &= \frac{1 + 2\sum_{i=1}^{n-1}\rho_i}{nh_n^p}\|K\|_2^2 \to 0,\end{aligned}$$

as $n \to \infty$, because of (3.iii). □

Lemma 3.2. *For $g_1, g_2 \geqslant 0$, with compact support in $[-1,1]^p$ there exist constants C_2, C_3 and C_4 such that*

$$|\bar{J}(g_2) - \bar{J}(g_1)| \leqslant C_2\sqrt{\bar{J}(g_1)}\|g_2 - g_1\|_2 + C_3\|g_2 - g_1\|_2^2 + C_4\|g_2 - g_1\|_2.$$

Proof. Write

$$|H(y) - H(x)| \leqslant [H'(x+) + H'(y+)]|y - x|,$$

giving

$$\begin{aligned}|\bar{J}(g_2) - \bar{J}(g_1)| &\leqslant \int [H'(g_2(x)+) + H'(g_1(x)+)]|g_2(x) - g_1(x)|\,\mathrm{d}x \\ &\leqslant \left(\sqrt{\int H'(g_2(x)+)^2\,\mathrm{d}x} + \sqrt{\int H'(g_1(x)+)^2\,\mathrm{d}x}\right)\|g_2 - g_1\|_2. \qquad (1)\end{aligned}$$

Condition (3.v) gives

$$H(2x)\geqslant H(x)+xH'(x+)\geqslant xH'(x+), \tag{2}$$

i.e.

$$\frac{H'(x+)}{x}\leqslant\frac{H(2x)}{x^2}\leqslant 4C_0, \tag{3}$$

when $x\geqslant 1$ by (3.vi). This entails

$$\begin{aligned} H'(x+)^2 &\leqslant 4C_0H'(x+)x\chi_{\{x\geqslant 1\}}+H'(1+)^2\chi_{\{x\leqslant 1\}} \\ &\leqslant 4C_0H(2x)+H'(1+)^2 \\ &\leqslant 4C_0C_1H(x)+H'(1+)^2, \end{aligned} \tag{4}$$

where the last two inequalities follow from (2) and (3.vii). Now, as $\operatorname{supp}(g_1)\subseteq[-1,1]^p$,

$$\begin{aligned} \sqrt{\int H'(g_1(x)+)^2\,\mathrm{d}x} &\leqslant \sqrt{4C_0C_1\int H(g_1(x))\,\mathrm{d}x+2^pH'(1+)^2} \\ &\leqslant \sqrt{4C_0C_1}\sqrt{\bar{J}(g_1)}+2^{p/2}H'(1+). \end{aligned} \tag{5}$$

Additionally,

$$\begin{aligned} H(g_2(x)) &\leqslant C_1(H(g_1(x))+H(|g_2(x)-g_1(x)|)) \\ &\leqslant C_1(H(g_1(x))+C_0C_1(g_2(x)-g_1(x))^2+C_1H(1), \end{aligned}$$

using (3.vii). That is,

$$\begin{aligned} \sqrt{\int H'(g_2(x)+)^2\,\mathrm{d}x} &\leqslant \sqrt{4C_0C_1}\sqrt{\bar{J}(g_2)}+2^{p/2}H'(1+) \\ &\leqslant \sqrt{4C_0C_1}\sqrt{C_1\bar{J}(g_1)+C_0C_1\|g_2-g_1\|_2^2+2^pC_1H(1)}+2^{p/2}H'(1+) \\ &\leqslant \sqrt{4C_0C_1^2}\sqrt{\bar{J}(g_1)}+2C_0C_1\|g_2-g_1\|_2+\sqrt{C_0C_1^22^{p+2}H(1)}+2^{p/2}H'(1+). \end{aligned} \tag{6}$$

Using (1), (5) and (6) we get the theorem, with $C_2=\sqrt{4C_0C_1}+\sqrt{4C_0C_1^2}$, $C_3=2C_0C_1$ and $C_4=2^{p/2+1}H'(1+)+\sqrt{C_0C_1^22^{p+2}H(1)}$. □

Lemma 3.3. $\bar{J}(\bar{f}_n)\leqslant\bar{J}(f)$.

Proof. Write

$$\begin{aligned} \bar{J}(\bar{f}_n) &= \int H\left(\frac{1}{h_n^p}\int K\left(\frac{x-y}{h_n}\right)\mathrm{d}y\right)\mathrm{d}x \\ &\leqslant \int\int \frac{1}{h_n^p}K\left(\frac{x-y}{h_n}\right)H(f(y))\,\mathrm{d}y\,\mathrm{d}x \\ &= \int H(f(y))\,\mathrm{d}y=\bar{J}(f), \end{aligned} \tag{7}$$

using Jensen's inequality and the convexity of H. □

4. Testing

Let $\mathscr{F}$ be the collection of probability measures on $\mathbb{R}^p$ with compact support. In order to test whether the marginal distribution of μ of $\{X_i\}$ has a singular part or not, introduce

$$\mathscr{F}_s = \{\mu \in \mathscr{F};\ \mu \text{ has a singular part}\},$$

$$\mathscr{F}_a = \{\mu \in \mathscr{F};\ \mu \ll \lambda\},$$

as a disjoint decomposition of $\mathscr{F}$. It is impossible to distinguish $\mathscr{F}_s$ from $\mathscr{F}_a$, since there are elements μ on the border between $\mathscr{F}_s$ and $\mathscr{F}_a$ (cf. Donoho, 1988). However, if we define ($f = \mathrm{d}\mu/\mathrm{d}\lambda$)

$$\mathscr{F}_a^{H,C} = \left\{\mu \in \mathscr{F}_a;\ \int H(f(x))\,\mathrm{d}x \leqslant C\right\},$$

we get a subset of $\mathscr{F}_a$ that can be asymptotically distinguished from $\mathscr{F}_s$. The testing procedure can be formalized as

$$\mathrm{H}_0:\ \mu \in \mathscr{F}_a^{H,C}$$

and

$$\mathrm{H}_1:\ \mu \in \mathscr{F}_s.$$

Then introduce a sequence of decision rules $\phi = \{\phi_n\}$ where $\phi_n : \mathbb{R}^{np} \to [0,1]$ and define the power function

$$\beta_n(\mu, \phi_n) = \mathrm{E}\phi_n(X_1, \dots, X_n),$$

and the asymptotic power function

$$\beta(\mu, \phi) = \lim_{n\to\infty} \beta_n(\mu, \phi_n).$$

Our test will be based on the theory of Sections 2 and 3. Given H and a sequence of thresholds $t = \{t_n\}_{n=1}^{\infty}$, put

$$\phi_n(X_1, \dots, X_n) = \chi_{\{\int H(\hat{f}_n(x))\,\mathrm{d}x > t_n\}}.$$

The corresponding sequence of decision will be denoted $\phi_{H,t}$. The main testing result may be formulated as:

Corollary 4.1. *Let $\{X_n\}_1^{\infty}$ be a stationary stochastic process with marginal distribution μ such that* (2.i) *is satisfied if $\mu \in \mathscr{F}_s$ and* (3.i) *if $\mu \in \mathscr{F}_a$. Given any function H satisfying* (2.iii), (3.iv)–(3.vii) *and a constant $C > 0$ we get*

$$\beta(\mu; \phi_{H,t}) = \begin{cases} 1 & \text{if } \mu \in \mathscr{F}_s, \\ 0 & \text{if } \mu \in \mathscr{F}_a^{H,C}, \end{cases}$$

provided $C < \liminf t_n \leqslant \limsup t_n < \infty$, the kernel K satisfies (3.ii) *and the sequence of bandwidths* (3.iii).

Proof. If $\mu \in \mathscr{F}$, $\phi_n \to 1$ a.s. by Theorem 2.1. Since ϕ_n is a 0–1 variable, we obtain

$$1 \geqslant \beta_n(\mu; \phi_n) = P(\phi_n = 1) \nearrow 1,$$

which means $\beta(\mu;\phi_{H,t})=1$. Let $\underline{t}=\liminf_{n\to\infty} t_n$. If $\mu\in\mathscr{F}_a^{H,C}$ we get, for large enough n,

$$\beta_n(\mu,\phi_n)\leqslant P\left(|\bar{J}(\hat{f}_n)-\bar{J}(f)|>\frac{\underline{t}-C}{2}\right)$$
$$\leqslant\frac{E|\bar{J}(\hat{f}_n)-\bar{J}(f)|}{(\underline{t}-C)/2}\to 0$$

by Theorem 3.1 and Markov's Inequality. □

Typical examples of $\mathscr{F}_a^{H,C}$ are $L^{1+\varepsilon}$ balls ($\varepsilon>0$),

$$\left\{\mu\in\mathscr{F}_a;\ \int f(x)^{1+\varepsilon}\,\mathrm{d}x\leqslant C\right\}.$$

For small ε, the $L^{1+\varepsilon}$ balls are quite large. There exist however $\mu\in\mathscr{F}_a$ that do not belong to any $L^{1+\varepsilon}$ ball. The set

$$\left\{\mu\in\mathscr{F}_a;\ \int f(x)(\log f(x))_+\,\mathrm{d}x\leqslant C\right\},$$

is larger than any $L^{1+\varepsilon}$ ball, since $x(\log x)_+$ increases more slowly to infinity than $x^{1+\varepsilon}$. In fact, the following proposition states that we may always find an H so that $\mu\in\mathscr{F}_a^{H,C}$:

Proposition 4.1. *Given $\mu\in\mathscr{F}_a$, there exists a function H satisfying the conditions of Corollary 4.1 and a constant $C>0$ such that $\mu\in\mathscr{F}_a^{H,C}$.*

Proof. Choose any version of $f\in\mathrm{d}\mu/\mathrm{d}\lambda$ and put $A_n=\{x\in\mathbb{R}^p;\ n-1\leqslant f(x)\leqslant n\}$, $n\in\mathbb{N}$. Then

$$1=\int f(x)\,\mathrm{d}x\geqslant\sum_{n=1}^{\infty}(n-1)\lambda(A_n)$$

so that with $\xi_n=n\lambda(A_n)$,

$$S=\sum_{1}^{\infty}\xi_n\leqslant\lambda(A_1)+2\sum_{1}^{\infty}(n-1)\lambda(A_n)\leqslant 1+2\cdot 1=3.$$

Pick an increasing sequence $\alpha_n\nearrow\infty$ such that $C=\sum_1^\infty\xi_n\alpha_n<\infty$. (This is always possible. Put, for example, $\alpha_k=n$ if $S(1-2^{-(n-1)})\leqslant\sum_{j=1}^{k-1}\xi_j<S(1-2^{-n})$. Then $\sum_1^\infty\xi_n\alpha_n\leqslant\sum_1^\infty n(S2^{-(n-1)})<\infty$.) Then construct H, first at the natural numbers through

$$H(n)=\sum_{k=0}^{n}\alpha_k,$$

with $\alpha_0=0$ and then through linear interpolation, so that H becomes piecewise linear. This makes H non-decreasing and, thus,

$$\int H(f(x))\,\mathrm{d}x\leqslant\sum_{n=1}^{\infty}H(n)\lambda(A_n)=\sum_{n=1}^{n}\left(\frac{1}{n}\sum_{k=1}^{n}\alpha_k\right)n\lambda(A_n)\leqslant\sum_{n=1}^{\infty}\alpha_n n\lambda(A_n)=C,$$

where the last inequality follows since $\{\alpha_n\}$ is increasing. It remains to verify all regularity conditions of H. Assume $x \in (n-1, n]$. Then

$$\frac{H(x)}{x} = \frac{\sum_{k=1}^{n-1} \alpha_k + (x-n+1)\alpha_n}{n-1+(x-n+1)} \geqslant \frac{1}{n-1}\sum_{k=1}^{n-1} \alpha_k \nearrow \infty,$$

as x and $n \nearrow \infty$, which proves (2.iii). Condition (3.v) also follows since $\{\alpha_n\}$ is increasing. Condition (3.iv) is obvious and (3.vi)–(3.vii) follow if we let $\alpha_n \nearrow \infty$ no faster than linearly. □

A limitation of Corollary 4.1 is that we cannot guarantee that $\beta_n(\cdot;\phi_n)$ converge uniformly to 1 over $\mathscr{F}_{\mathrm{s}}$ or to 0 over $\mathscr{F}_{\mathrm{a}}^{H,C}$. To get a practically more useful decision rule we restrict ourselves to independent data and choose a subset of $\mathscr{F}_{\mathrm{s}}$. Loosely speaking, the singular part of μ has to be large enough in order to distinguish it from $\mathscr{F}_{\mathrm{a}}^{H,C}$ for a fixed n.

One possibility is: Test H_0 against

$$\mathrm{H}_1': \mu \in \mathscr{F}_{\mathrm{s}}^{\delta,\alpha} = \{\mu \in \mathscr{F}_{\mathrm{s}};\ \exists B \text{ such that } \mu(B) \geqslant \delta,\ \lambda(B^h) \leqslant \alpha(h),\ \forall h > 0\},$$

with $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$, $\alpha(h) \searrow 0$ as $h \searrow 0$. If for example $\delta = 1$, $\alpha(h) = Ah^d$ and $0 < d \leqslant p$, then $\mu \in \mathscr{F}_{\mathrm{s}}^{\delta,\alpha}$ means that supp(μ) has Box dimension $\leqslant p-d$ (cf. Frigyesi and Hossjer, 1996).

Corollary 4.2. *Assume the same conditions as in Corollary* 4.1, *except that* $\{X_n\}_1^\infty$ *is a sequence of independent and identically distributed random variables. Then*

$$\inf_{\mu \in \mathscr{F}_{\mathrm{s}}^{\delta,\alpha}} \beta_n(\mu, \phi_n) \geqslant \bar{\beta}_n \nearrow 1,$$

$$\sup_{\mu \in \mathscr{F}_{\mathrm{a}}^{H,C}} \beta_n(\mu, \phi_n) \leqslant \underline{\beta}_n \searrow 0$$

as $n \to \infty$, *where explicit formulas for* $\bar{\beta}_n$ *and* $\underline{\beta}_n$ *are given below.*

Proof. Note that, by the proof of Theorem 2.1,

$$\bar{J}(\hat{f}_n) \geqslant \underline{H}\left(\frac{N/n}{2\alpha(h_n)}\right)\frac{N/n}{2},$$

if supp(K) $\subseteq [-1,1]^p$ (otherwise, the lower bound is slightly different). Since we have independent observations,

$$N \in Bin(n, \mu(B)) \overset{\mathrm{s}}{>>} N_\delta \in Bin(n, \delta),$$

where $\overset{\mathrm{s}}{>>}$ means "stochastically larger than". Thus

$$\inf_{\mu \in \mathscr{F}_{\mathrm{s}}^{\delta,\alpha}} \beta_n(\mu, \phi_n) \geqslant P\left(\underline{H}\left(\frac{N_\delta/n}{2\alpha(h_n)}\right)\frac{N_\delta}{n} > 2t_n\right) = \bar{\beta}_n \to 1,$$

when $n \to \infty$, as $N_\delta/n \xrightarrow{p} \delta$, $h_n \searrow 0$ and $\underline{H}(x) \nearrow \infty$ as $x \nearrow \infty$. If $\mu \in \mathscr{F}_{\mathrm{a}}^{H,C}$, we get from the proof of Theorem 3.1 and Lemma 3.1,

$$\begin{aligned} \mathrm{E}|\bar{J}(\hat{f}_n) - \bar{J}(\bar{f}_n)| &\leqslant (C_2\sqrt{C} + C_4)(\mathrm{E}\|\hat{f}_n - \bar{f}_n\|^2)^{1/2} + C_3\mathrm{E}\|\hat{f}_1 - \bar{f}_n\|^2 \\ &\leqslant \frac{(C_2\sqrt{C} + C_4)\|K\|_2}{n^{1/2}h_n^{p/2}} + \frac{C_3\|K\|_2^2}{nh_n^p}, \end{aligned}$$

using the fact that $\bar{J}(f) \leqslant C$. Since $\bar{J}(\bar{f}_n) \leqslant \bar{J}(f) \leqslant C$ (Lemma 3.3), Markov's inequality yields, for all n large enough,

$$\sup_{\mu \in \mathscr{F}_s^{\delta,\alpha}} \beta_n(\mu, \phi_n) \leqslant \sup_{\mu \in \mathscr{F}_s^{\delta,\alpha}} \frac{\mathrm{E}|\bar{J}(\hat{f}_n) - \bar{J}(\bar{f}_n)|}{(\bar{t} - C)/2}$$

$$\leqslant \frac{2(C_2\sqrt{C} + C_4)\|K\|_2}{n^{1/2}h_n^{p/2}(\bar{t} - C)} + \frac{2C_3\|K\|_2^2}{nh_n^p(\bar{t} - C)} = \underline{\beta}_n \searrow 0,$$

as $n \to \infty$. □

Remark. We may choose $t_n \searrow C$ with such a speed that $\underline{\beta}_n = \alpha_0$ for some fixed level α_0, replacing $(\bar{t} - C)/2$ by $t_n - C$.

5. Conclusion

We have seen how certain functionals, notably the L^p-norm, of kernel density estimators can be used to test whether data come from a singular or an absolutely continuous distribution function. The functionals diverge as the sample size is increased in the singular case, and converge for a wide class of absolutely continuous distributions.

This work resulted in another paper by the authors, where the velocity of divergence of $J(\hat{f}_n)$ was investigated when $H(x) = x^{1+\varepsilon}$. It turned out that this velocity depends on the geometrical properties of the measure, namely the εth generalized Rényi dimension. These ideas, and how to use them to estimate fractal dimensions are investigated in Frigyesi and Hössjer (1996).

Acknowledgements

We would hereby like to express our gratitude towards Professor Nils Nilsson for his kind assistance with Lemma 2.1. We are also grateful to a referee, who urged us to further develop the testing idea of Section 4.

References

Cohn, D.L., 1980. Measure Theory. Birkhäuser, Boston.
Donoho, D.L., 1988. One-sided inference about functionals of a density. Ann. Statist. 10, 1390–1420.
Frigyesi, A., Hössjer, O., 1996. Kernel estimates of dimension spectra for multifractal measures. Technical Report at the Department of Mathematical Statistics at Lund University, Sweden.
Hall, P. Heyde, C.C., 1980. Martingale Limit Theory and Its Applications. Academic Press, New York.
Wand, M.P., Jones, M.C., 1995. Kernel Smoothing. Chapman & Hall, London.