Bernoulli **7**(3), 2001, 507–525

Minimax hypothesis testing about the density support

GHISLAINE GAYRAUD

Laboratoire de Statistique, CREST, Timbre J340, 3 avenue Pierre Larousse, 92241 Malakoff Cedex, France; and Laboratoire de Mathématiques Raphael Salem, UMR CNRS 6085, UFR des Sciences-Mathématiques, 76821 Mont St Aignan Cedex, France. E-mail: gayraud@ensae.fr

The paper is concerned with testing nonparametric hypotheses about the underlying support G of independent and identically distributed observations. It is assumed that G belongs to a class $\mathscr{G}$ of compact sets with smooth upper surface called boundary fragments. It is required to distinguish the simple null hypothesis specified by a known set G_0 in $\mathscr{G}$ against nonparametric alternatives that G belongs to a class obtained by removing a certain neighbourhood of G_0 in $\mathscr{G}$. Using the asymptotic minimax approach, the problem is to determine the order of the smallest distance between the null hypothesis H_0 and the alternatives for which one is able to test the null hypothesis against the alternatives with a given summarized error.

Keywords: minimax rate of testing; nonparametric hypothesis testing; underlying support

1. Introduction

We observe N-dimensional independent and identically distributed random variables $X_1, \ldots, X_n$, uniformly distributed on an unknown set G. We assume that leb(G), the Lebesgue measure of G, is positive. We denote by $\Sigma_{N-1}(\gamma, L_1)$ the class of functions on $[0, 1]^{N-1}$ having continuous partial derivatives up to order $k = \lfloor \gamma \rfloor$ ($k \in \mathbb{N}$ is the greater integer strictly less than γ) and such that

$$|g(z) - p_y^g(z)| \leqslant L_1 |z - y|^\gamma, \qquad \forall y, z \in [0, 1]^{N-1},$$

where $p_y^g(z)$ denotes the Taylor polynomial of order k for $g(\cdot)$ at a point y, and $|z|$ denotes the Euclidean norm of a vector z. We assume that the set $G \subset [0, 1]^N$ is of the form $G = \{x = (x_1, \ldots, x_N) \in [0, 1]^N : 0 \leqslant x_N \leqslant g(x_1, \ldots, x_{N-1})\}$, where $g : [0, 1]^{N-1} \rightarrow [0, 1]$ is called the *edge* of G (Korostelev and Tsybakov 1993a) and is a smooth function belonging to $\Sigma(\gamma, L_1, b_1)$, which is defined by

$$\Sigma(\gamma, L_1, b_1) = \{g \in \Sigma_{N-1}(\gamma, L_1) : b_1 < g(y) < 1 - b_1 \ \forall y \in [0, 1]^{N-1}\},$$

where $\gamma \geqslant 1$ is real, L_1 is a positive constant and b_1 is a positive constant such that $0 < b_1 < \frac{1}{2}$. We denote by $\mathscr{G}$ the class of sets whose edge belongs to $\Sigma(\gamma, L_1, b_1)$. Such sets G are called *boundary fragments* (Korostelev and Tsybakov 1993b).

In this framework, the present paper studies the following hypothesis testing problem concerning the support G: the null hypothesis is specified by a known fixed set G_0 in $\mathscr{G}$ and the alternatives are classes of sets, obtained by removing a certain ψ_n-neighbourhood of G_0 in $\mathscr{G}$, where ψ_n is a sequence of positive numbers decreasing to zero with n. In detail, we first let d_∞ be the Hausdorff distance between two closed compact sets G and G' defined by $d_\infty(G, G') = \max(\max_{x\in G}\rho(x, G'), \max_{x\in G'}\rho(x, G))$, where $\rho(x, G)$ is the Euclidean distance between a point x and a closed set G. We consider the problem of testing the simple hypothesis

$$H_0 : G = G_0,$$

against the composite alternative

$$H_{\Lambda_{1,n}} : G \in \Lambda_{1,n}(\psi_n) = \{G \in \mathscr{G} : d_\infty(G_0, G) \geqslant \psi_n\}.$$

Since $\gamma \geqslant 1$, the Hausdorff distance d_∞ between G and G_0 is equivalent to the L_∞-distance between the corresponding edge functions g and g_0. Thus, $\Lambda_{1,n}(\psi_n)$ can be defined as the class of sets whose edge belongs to $\Sigma(\gamma, L_1, b_1)$ and is separated from g_0 in L_∞-distance by $c\psi_n$, where c is a positive constant.

Second, consider support functionals $S(G)$ defined by $S(G) = \int_G \varphi(x)\mathrm{d}x$, where φ is some known bounded positive function on $[0, 1]^N$; also let $S_0 = S(G_0)$. The test we are interested in is

$$H_0 : G = G_0,$$

against the composite alternative

$$H_{\Lambda_{2,n}} : G \in \Lambda_{2,n}(\psi_n) = \{G \in \mathscr{G} : |S(G) - S_0| \geqslant \psi_n\}.$$

Third, let $d_1(G, G')$ be the Lebesgue measure of the symmetric difference between two compact closed sets G and G'. We wish to test

$$H_0 : G = G_0,$$

against the composite alternative

$$H_{\Lambda_{3,n}} : G \in \Lambda_{3,n}(\psi_n) = \{G \in \mathscr{G} : d_1(G, G_0) \geqslant \psi_n\}.$$

Since the Lebesgue measure of the symmetric difference d_1 between G and G_0 is equal to the L_1-distance between the corresponding edge functions g and g_0, $\Lambda_{3,n}(\psi_n)$ can be defined as the class of sets in $\mathscr{G}$ whose edge is separated from g_0 in L_1-distance by ψ_n.

$\Lambda_{1,n}$, $\Lambda_{2,n}$ and $\Lambda_{3,n}$ are henceforth abbreviated as Λ_n when it is convenient, and **d** represents the distance which separates the alternatives from G_0. Note that Λ_n is defined by three parameters: the class $\mathscr{G}$, **d** and ψ_n. However, it can be shown (Ingster 1993a; 1993b; 1993c) that, given $\mathscr{G}$ and **d**, ψ_n cannot be chosen arbitrarily. It turns out that if ψ_n is too small, then it is not possible to test the hypothesis H_0 against H_{Λ_n} with given summarized errors of the first and the second kind. On the other hand, if ψ_n is very large, such a test is possible; the problem is to find the smallest ψ_n for which such a test is still possible and to indicate the corresponding test.

Let us give some precise definitions to solve these problems. Let Δ_n be a *test statistic* that is an arbitrary function with values 0, 1 and is measurable with respect to $X_1, \ldots, X_n$ and such that we accept H_0 if $\Delta_n = 0$ and we reject H_0 if $\Delta_n = 1$. Set $R_0(\Delta_n) = P_{G_0}\{\Delta_n = 1\}$ for the error of the first kind and $R_1(\Delta_n, \psi_n) = \sup_{G \in \Lambda_n(\psi_n)} P_G\{\Delta_n = 0\}$ for the error of second kind. The index G means that the measure P_G is generated by $X_1, \ldots, X_n$ uniformly distributed on G. The properties of the tests Δ_n are characterized by the sum of the two errors

$$R_0(\Delta_n) + R_1(\Delta_n, \psi_n).$$

Fix a number β in (0, 1). The sequence ψ_n is called the minimax rate of testing (MRT) if the following two conditions hold:

- there exists $b > 0$ such that

$$\liminf_{n\to\infty} \inf_{\Delta_n} [R_0(\Delta_n) + R_1(\Delta_n, b\psi_n)] \geqslant \beta, \tag{1.1}$$

where the infimum is taken over all the tests Δ_n;

- there exist a positive constant b' and a test $\overline{\Delta}_n$ such that

$$\limsup_{n\to\infty} [R_0(\overline{\Delta}_n) + R_1(\overline{\Delta}_n, b'\psi_n)] \leqslant \beta. \tag{1.2}$$

Thus, the MRT ψ_n is such that a meaningful test of H_0 is impossible if the distance between the null hypothesis and the alternative is smaller than $b\psi_n$, and is possible if the distance is greater than $b'\psi_n$. Clearly, $b' \geqslant b$.

The problem of nonparametric hypothesis testing is closely related to minimax nonparametric estimation problems with infinite-dimensional 'parameter' class. For a detailed review of this area, see Donoho *et al.* (1995) and the references therein. However, specific features of hypothesis testing problems have been discovered which do not occur in estimation problems. In particular, the minimax rates are different for the estimation problem and for the testing problem. The problem of nonparametric hypothesis testing was initiated by Ingster (1982), although some closely related ideas appeared in earlier papers by Burnashev (1979), Ibragimov and Has'minskii (1977) and Birgé(1983). Ingster (1982) considered hypothesis testing in the problem of signal detection where a certain function f, treated as a signal, is observed with noise: the null hypothesis in this case is $f \equiv 0$, that is, no signal is present. The problem of testing this null hypothesis against the alternative defined as the set of functions belonging to an ellipsoid in $L_2(0, 1)$ and separated from zero in the L_2-norm by a distance $K\psi_n$ (K is a positive constant) was studied by Ingster (1982) and Ermakov (1990a). Another alternative defined as a Hölder class separated from zero in the L_p-norm, in the uniform norm and in a fixed point by a distance $K\psi_n$ was considered by Ingster (1986b). In addition, Ingster (1990) and Suslina (1993) investigated the case of an alternative defined as a set of functions belonging to an ellipsoid in l_p, $0 < p \leqslant \infty$, lying outside the ball of radius $K\psi_n$ around zero. Lepski (1993) studied a slightly different problem: that of finding the exact value $K\psi_n$, where K is a positive constant depending on the smoothness parameter, for which relations (1.2) and (1.1) hold with alternative sets defined by the Hölder class separated from zero in the uniform norm and also in a fixed point; an extension of Lepski's (1993) result is given in Lepski and Tsybakov (1996). Lepski and Spokoiny (1999) studied the signal detection problem, considering a Besov ball as the alternative separated away from zero in the integral L_p-norm. Spokoiny (1996) extended the investigations to the problem of adaptive testing. This problem of adaptive testing is also studied in Spokoiny (1998). Another widely studied example of hypothesis testing concerns the probability density; this can be formulated in the following way. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having an unknown probability density f. The null hypothesis is specified by a known density f_0 against several nonparametric alternatives such as ellipsoids in L_2 (Ingster 1984; Ermakov 1994), Hölder classes (Ingster 1986b), ellipsoids in l_p, $0 < p \leqslant \infty$ (Ingster 1994), and Sobolev balls (Ingster 1986a). The reader is referred to Ingster (1993a; 1993b; 1993c) for a most detailed review of nonparametric minimax hypothesis testing for both signal detection and density problems.

Other problems of nonparametric hypothesis testing are studied using the minimax approach: see, for instance, Ermakov (1990b; 1990c) in which the objects of interest are respectively the spectral density and the distribution function, and recent papers (Ermakov 1996; Härdle *et al.* 1997; Spokoiny 1997; Feldmann *et al.* 1998; Baraud *et al.* 1999; Gayraud and Tsybakov 1999; Härdle and Kneip 1999; Pouet 2000) in which the nonparametric null hypothesis and alternatives are both composite.

Although our problem is a minimax hypothesis testing problem, it differs from those mentioned above in that it concerns a set (the underlying support) and not a function. As in signal detection (Ingster 1982), we show that the minimax rates are sometimes different for the estimation problem and for the hypothesis testing problem: actually the MRT is at least the same as the minimax rate of estimation (MRE), or else the MRT improves the MRE obtained in the corresponding estimation problems. Our study not only is of theoretical interest but also could be important in many applications. Indeed, the underlying support is an object of interest in several areas such as econometrics, cluster analysis and reliability theory. For example, the knowledge of the boundary of the density support allows the performance of an enterprise to be evaluated in terms of technical efficiency, and also the underlying support can be a useful tool in reliability theory for detecting abnormal system behaviour.

The paper is organized as follows. In Section 2, we present the test statistics for which relation (1.2) holds. The main results of this paper are stated in Section 3 and their proofs are given in Section 5. We prove that the MRT is either $(n/\log n)^{-\gamma/(\gamma+N-1)}$ for the alternative $\Lambda_{1,n}(\psi_n)$, or $n^{-[\gamma+(N-1)/2]/(\gamma+N-1)}$ when the alternative is defined by either $\Lambda_{2,n}(\psi_n)$ or $\Lambda_{3,n}(\psi_n)$. A comparison between the MRT and the MRE obtained in the corresponding estimation problems shows that they are equal when both alternative and error of estimation are defined by the d_∞-distance and also by the positive difference $|S(G) - S_0|$. When both alternative and estimation error are defined with the d_1 metric, it is interesting to note that the MRT and the MRE are different and, in particular, that the MRT improves the MRE. In this last case, the MRT corresponds to the MRE obtained in the estimation problem of functionals of support such as $T(G) = \int_{[0,1]^{N-1}} |g(y) - g_0(y)|\mathrm{d}y$ (Gayraud 1997). Section 4 is devoted to additional remarks and simulations.

2. Definition of the test statistics

Henceforth, let δ_n be a positive sequence and set $M = \delta_n^{-(N-1)}$. Without loss of generality, suppose that M is an integer. Introduce a partition of $[0, 1]^{N-1}$ into cubes Q_q, $q = 1, \ldots, M$, with edges of length δ_n. For each q in $\{1, \ldots, M\}$, denote by $u_q = (u_{q,1}, \ldots, u_{q,N-1}) \in [0, 1]^{N-1}$ the centre of the cube Q_q.

2.1. Testing for supports in the Hausdorff distance

In this section, set $\delta_n = (n/\log n)^{-(1/\gamma+N-1)}$. The test statistic is based on G_n and g_n, the estimates of G and its edge g proposed in Korostelev and Tsybakov (1993b); there, edge estimation is carried out separately on each cube Q_q, $q \in \{1, \ldots, M\}$, as a polynomial function, and thus the entire process of estimating g_n is based on slicing and piecewise polynomial approximation; then, the set estimator G_n is defined as the set

$$G_n = \{x = (y, x_N) \in [0, 1]^N : 0 \leqslant x_N \leqslant g_n(y)\}. \tag{2.1}$$

This leads to the test statistic

$$\overline{\Delta}_{1,n} = \begin{cases} 0 & \text{if } d_\infty(G_n, G_0) < C_1\delta_n^\gamma, \\ 1 & \text{if } d_\infty(G_n, G_0) \geqslant C_1\delta_n^\gamma, \end{cases}$$

where $C_1 > (N-1)/(\gamma + N - 1)$ is a constant.

2.2. Tests for functionals of density supports

In this section, set $\delta_n = n^{-(1/\gamma+N-1)}$. This problem is related to the minimax estimation of the functionals $S(G)$ (Gayraud 1997). We first define an estimator S_n of $S(G)$: divide the whole sample $X_1, \ldots, X_n$ into three subsamples $J_1 = \{X_i, i \in I_1\}$, $J_2 = \{X_i, i \in I_2\}$, $J_3 = \{X_i, i \in I_3\}$ such that $I_1 \cup I_2 \cup I_3 = \{1, \ldots, n\}$ and $\mathrm{card}\, J_1 = \mathrm{card}\, J_2 = \mathrm{card}\, J_3 = n/3$. Without loss of generality, suppose that $n/3$ is an integer. We transform the estimator G_n defined in (2.1) as follows: instead of using the original sample, we consider another sample $X'_1, \ldots, X'_n$, obtained by a transformation Υ of $X_1, \ldots, X_n$; this allows us to construct an estimator G_n included in the true support G almost surely (the proof and the transformation Υ are given in Gayraud 1997). Then, let J'_1 and J'_3 be the samples obtained by transformation Υ of J_1 and J_3, and denote by G_{n,J'_r}, g_{n,J'_r} and $\overline{G}_{n,J'_r}$ the estimator of G, the estimator of g, and the complement to G_{n,J'_r} in $[0, 1]^N$, respectively; all of these are based on J'_r, for all $r \in \{1, 3\}$. The estimator S_n of $S(G)$ is defined by

$$S_n = \int_{[0,1]^N} \varphi(x) I\{x_N \leqslant g_{n,J'_1}(x_1, \ldots, x_{N-1})\} \mathrm{d}x + \frac{3}{n} \sum_{i \in I_2} \varphi(X_i) I\{X_i \in \overline{G}_{n,J'_1}\} \mu_{n,J'_3}, \quad (2.2)$$

where μ_{n,J'_3}, which is based on J'_3, is the estimator of leb(G) defined in Korostelev and Tsybakov (1993b, Lemma 4). We can now define the test statistic

$$\overline{\Delta}_{2,n} = \begin{cases} 0 & \text{if } (S_n - S_0)^2 < C_2 \delta_n^\gamma / n, \\ 1 & \text{if } (S_n - S_0)^2 \geqslant C_2 \delta_n^\gamma / n, \end{cases}$$

where C_2 is a positive constant.

2.3. Tests for supports in the d_1-distance

In this section, set $\delta_n = n^{-(1/\gamma + N - 1)}$. Although this problem is related to the density support estimation, the construction of the test is based on an estimator T_n of $T(G) = \int_{[0,1]^{N-1}} |g(y) - g_0(y)| \mathrm{d}y$. We first divide the whole sample $X_1, \ldots, X_n$ into two subsamples J_1, J_2 and then divide each subsample J_q into three sub-subsamples $J_{q,1}$, $J_{q,2}$, $J_{q,3}$, $q = 1, 2$. Without loss of generality, we suppose that $n/6$ is an integer and that each sub-subsample $J_{q,1}$, $J_{q,2}$, $J_{q,3}$, $q = 1, 2$, contains $n/6$ observations. Moreover, define $J'_{q,r}$, $q = 1, 2$ and $r \in \{1, 3\}$, the sample obtained by the transformation Υ of $J_{q,r}$ as in Section 2.2. The functional $T(G)$ can be written as the sum of two terms:

$$T(G) = \int (g(y) - g_0(y)) I\{g(y) \geqslant g_0(y)\} \mathrm{d}y + \int (g_0(y) - g(y)) I\{g(y) < g_0(y)\} \mathrm{d}y = t_1 + t_2.$$

Then define the statistic

$$T_n = \hat{t}_1 + \hat{t}_2, \quad (2.3)$$

where

$$\hat{t}_1 = \int_{[0,1]^N} I\{x \in G_{n,J'_{1,1}} \cap \overline{G}_0\} \mathrm{d}x + \frac{6}{n} \sum_{i:X_i \in J_{1,2}} I\{X_i \in \overline{G}_{n,J'_{1,1}} \cap \overline{G}_0\} \mu_{n,J'_{1,3}},$$

$$\hat{t}_2 = \int_{[0,1]^N} I\{x \in \overline{G}_{n,J'_{2,1}} \cap G_0\} \mathrm{d}x - \frac{6}{n} \sum_{i:X_i \in J_{2,2}} I\{X_i \in \overline{G}_{n,J'_{2,1}} \cap G_0\} \mu_{n,J'_{2,3}},$$

where $G_{n,J'_{q,1}}$ is the set estimator defined as in Section 2.2, which is based on $J'_{q,1}$ for $q = 1, 2$, and $\mu_{n,J'_{q,3}}$ is the leb(G)-estimator defined in Korostelev and Tsybakov (1993b, Lemma 4) and based on $J'_{q,3}$ for $q = 1, 2$. Then, our test statistic is

$$\overline{\Delta}_{3,n} = \begin{cases} 0 & \text{if } |T_n| < C_3 (\delta_n^\gamma / n)^{1/2}, \\ 1 & \text{if } |T_n| \geqslant C_3 (\delta_n^\gamma / n)^{1/2}, \end{cases}$$

where C_3 is a positive constant.

3. Main results

Theorem 3.1. *Let alternatives be defined by the set* $\Lambda_{1,n}(\psi_n)$*, where* $\psi_n = (n/\log n)^{-(\gamma/\gamma+N-1)}$. *There exist positive constants* b_2 *and* b_3 *for which the following relations hold:*

$$\liminf_{n\to\infty} \inf_{\Delta_n} [R_0(\Delta_n) + R_1(\Delta_n, b_2\psi_n)] \geqslant 1, \tag{3.1}$$

$$\limsup_{n\to\infty} [R_0(\overline{\Delta}_{1,n}) + R_1(\overline{\Delta}_{1,n}, b_3\psi_n)] = 0, \tag{3.2}$$

where $\inf_{\Delta_n}$ *denotes the infimum over all test statistics.*

Remark 3.1. Note that ψ_n corresponds to the MRE obtained in Korostelev and Tsybakov (1993b) for density support estimation with the Hausdorff distance for the class of boundary fragments.

Remark 3.2. In Theorem 3.1, the right-hand side of each relation does not depend on β and therefore (3.1) and (3.2) are satisfied for any $\beta \in [0, 1]$. This is connected with the fact that the limiting distribution, arising here, is singular.

Theorem 3.2. *Let alternatives be defined by the set* $\Lambda_{2,n}(\psi_n)$*, where* $\psi_n = (\delta_n^\gamma/n)^{1/2}$ *and* $\delta_n = n^{-(1/\gamma+N-1)}$.

(*i*) *Assume that* φ *is an integrable function on* $[0, 1]^N$ *such that* $|\varphi|$ *is greater than a positive constant on some closed* N*-interval contained in* $[0, 1]^{N-1} \times [b_1, 1 - b_1]$. *Then, there exist positive constants* b_4 *and* $\beta^* < 1$ *such that, for all* $\beta < \beta^*$*, the following inequality holds:*

$$\liminf_{n\to\infty} \inf_{\Delta_n} [R_0(\Delta_n) + R_1(\Delta_n, b_4\psi_n)] \geqslant \beta, \tag{3.3}$$

where $\inf_{\Delta_n}$ *denotes the infimum over all possible test statistics.*

(*ii*) *Assume that* φ *is continuous and bounded on* $[0, 1]^N$. *Then, there exists some positive constant* b_5 *such that*

$$\limsup_{n\to\infty} [R_0(\overline{\Delta}_{2,n}) + R_1(\overline{\Delta}_{2,n}, b_5\psi_n)] \leqslant \beta. \tag{3.4}$$

Remark 3.3. Note that ψ_n is equal to the MRE obtained in Gayraud (1997) for the estimation of the functional density support for the class of boundary fragments.

Theorem 3.3. *Let alternatives be defined by the set* $\Lambda_{3,n}(\psi_n)$*, where* $\psi_n = (\delta_n^\gamma/n)^{1/2}$ *and* $\delta_n = n^{-(1/\gamma+N-1)}$.

(*i*) *There exist positive constants* b_6 *and* $\beta^* < 1$ *such that,* $\forall \beta < \beta^*$*, we have*

$$\liminf_{n\to\infty} \inf_{\Delta_n} [R_0(\Delta_n) + R_1(\Delta_n, b_6\psi_n)] \geqslant \beta, \tag{3.5}$$

where $\inf_{\Delta_n}$ *denotes the infimum over all possible test statistics.*

(*ii*) *There exists some positive constant* b_7 *such that*

$$\limsup_{n\to\infty}[R_0(\overline{\Delta}_{3,n}) + R_1(\overline{\Delta}_{3,n}, b_7\psi_n)] \leqslant \beta. \tag{3.6}$$

Remark 3.4. In this case, the MRT ψ_n improves the MRE obtained for the estimation of the density support when the error is defined with the d_1-metric (Gayraud 1997).

4. Additional remarks and simulations

4.1. Remarks

4.1.1. The case of an unknown probability density

The results of this paper can be generalized to a more general class of density than the uniform density. Let $\mathscr{F}_G$ be the class of densities whose underlying support is G such that

$$\mathscr{F}_G = \{f \in \tilde{\mathscr{F}}(a_0, G) : f \text{ has continuous partial derivatives up to order}$$
$$l-1 \text{ in } \mathrm{Int}(G) \text{ and } |f(x) - p_v^f(x)| \leqslant Q_L|x-v|^l,\ \forall x \in G,\ \forall v \in \mathrm{Int}(G)\},$$

where $p_v^f(x)$ is the Taylor polynomial of f of order $l-1$ at the point $v \in \mathrm{Int}(G)$, Q_L is a positive constant, $\mathrm{Int}(G)$ denotes the interior of G, l is a positive integer and the class $\tilde{\mathscr{F}}(a_0, G) = \{f$ defined on $[0, 1]^{\mathrm{N}} : f(x) \geqslant a_0 > 0,\ \forall x \in G$, and $f(x) = 0,\ \forall x \notin G\}$, where $a_0 > 0$ is a given constant. In this case, one defines a density estimator as a kernel estimator K (for its construction, see Gayraud 1997, Section 2.3) in place of the estimator of $\mathrm{leb}(G)$ used in the definition of both S_n (2.2) and T_n (2.3). Some assumptions on K and l (for details, see Gayraud 1997, Section 3.1) allow one to consider the probability density as a nuisance parameter. Then one obtains Theorems 3.2 and 3.3, since Theorem 3.1 remains valid.

4.1.2. Lower bounds

In nonparametric estimation problems such as regression or density estimation, the minimax lower bounds lead to proof that the rates of convergence obtained for some estimators cannot be improved by any other estimators. In hypothesis testing problems, the relations of the lower bound (1.1) lead to proof that relation (1.2) cannot be used with $\psi'_n = o(\psi_n)$ in place of $b'\psi_n$, that is, one cannot successfully distinguish the null hypothesis from the alternatives that are much closer than ψ_n from H_0 in **d**-distance. The difficulty in proving the relations of lower bounds in hypothesis testing problems lies in the construction of the parametric family which must be included in the whole class $\mathscr{G}$ as in nonparametric estimation problems, but also which must be separated from the null hypothesis by ψ_n in **d**-distance. This is achieved by randomizing the alternative classes of sets.

4.1.3. The exact separation constant

A possible extension of our results would be to provide b, the exact separation constant (ESC), for which (1.1) holds for all $b'' < b$ and (1.2) holds for all $b'' > b$. The ESC is known in several problems, in particular for functional classes and distances $\mathbf{d}$ defined in a coordinate form – ellipsoids in l_p in the density model in Ingster (1994) and in the signal detection problem in Suslina (1993). For the classes defined in functional form such as Hölder or Sobolev classes, with $\mathbf{d}$ defined by the L_p-norm, much less is known about the exact asymptotics: to our knowledge, the ESC is known only in the signal detection problem in Lepski (1993) for the Hölder class with smoothness parameter less than 1 and the L_∞-norm as the distance $\mathbf{d}$, and, in Lepski and Tsybakov (1996) and in Pouet (1999) for the Hölder and Sobolev classes and for analytical alternatives, respectively; in both papers $\mathbf{d}$ is defined by the supremum norm and by the distance in a fixed point.

In our framework, a study of the ESC would be a non-trivial matter requiring further investigations and requiring a paper in its own right.

4.2. Simulations

In this subsection we illustrate our theoretical results by comparing the errors of the second kind obtained under some alternative class of sets which are separated from H_0 by a, with the distance d_∞ and the distance based on the functional $S(\cdot)$. For this comparison, we consider the particular case of: $N = 2$; the null hypothesis $G_0 = \{x = (x_1, x_2) \in [0, 1]^2 : 0 \leqslant x_2 \leqslant g_0(x_1) = \frac{1}{2}\}$, $S(G) = \int_G \mathrm{d}x = \mathrm{leb}(G)$; distances $d_\infty(G, G_0)$, $|S(G) - S_0|$ used to separate G_0 and sets G belonging to the alternative; and two forms of alternative class, defined by

$$\mathscr{C}_1(a) = \{G : G = \{(x_1, x_2) : 0 \leqslant x_2 \leqslant g_0(x_1) + a\}\},$$

$$\mathscr{C}_2(a) = \{G : G = \{(x_1, x_2) : 0 \leqslant x_2 \leqslant g_0(x_1) + \kappa(a, \mathbf{d}) a \sin(x_1/a)\}\},$$

where a and $\kappa(a, \mathbf{d})$ are positive constants. The constant $\kappa(a, \mathbf{d})$ is chosen such that $\mathbf{d}(G, G_0) \geqslant a$, for all $G \in \mathscr{C}_2(a)$. Since the alternative hypotheses are composite, we takes a as varying inside a set $\mathscr{A}$ defined by $\mathscr{A} = \{0.03, 0.05, 0.07, 0.09, 0.1, 0.12\}$. Then we calculate $R_1^G(\overline{\Delta}_{1,n}, a) = P_G(\overline{\Delta}_{1,n} = 0)$ and $R_1^G(\overline{\Delta}_{2,n}, a) = P_G(\overline{\Delta}_{2,n} = 0)$, where G belongs to either $\mathscr{C}_1(a)$ or $\mathscr{C}_2(a)$ and a is in $\mathscr{A}$ (Table 1). The first step of these calculations is to compute the test statistics $\overline{\Delta}_{1,n}$ and $\overline{\Delta}_{2,n}$; this is done following the theoretical procedure given in Korostelev and Tsybakov (1993b) and in Gayraud (1997), respectively. The second step consists in using the Monte Carlo method with 10 000 replications to approximate each $R_1^G(\overline{\Delta}_{1,n}, a)$ and $R_1^G(\overline{\Delta}_{2,n}, a)$ for G in $\mathscr{C}_1(a) \cup \mathscr{C}_2(a)$ and a in $\mathscr{A}$. Since our theoretical results are asymptotic, our calculations are done with different values of n, that is, $n \in \{100, 250, 500, 750, 1000\}$. Furthermore, the value of $a = 0$ leads us to evaluate the error of the first kind since $R_1^G(\overline{\Delta}_{q,n}, 0) = 1 - R_0(\overline{\Delta}_{q,n})$, $q \in \{1, 2\}$. One must note that for each distance $\mathbf{d}$, the theoretical error of second kind defined in Section 1 is the maximal error over the class of sets G in $\mathscr{C}_1(a) \cup \mathscr{C}_2(a)$ and the set $\mathscr{A}$ of a. The presence of two different forms of alternative classes would demonstrate that the simulation results are independent of the choice of one particular form.

Table 1. Errors of first and second kind for G either in $\mathscr{C}_1(a)$ or in $\mathscr{C}_2(a)$

$R_1^G(\overline{\Delta}_{1,n}, a),\ G \in \mathscr{C}_1(a)$

Values of a	$n = 100$	$n = 250$	$n = 500$	$n = 750$	$n = 1000$
0	0.8598	0.8648	0.8877	0.8688	0.86
0.03	0.9581	0.9849	0.9952	0.0771	0.013
0.05	0.8798	0.0874	0	0	0
0.07	0.3652	0.0001	0	0	0
0.09	0.0112	0	0	0	0
0.1	0.0026	0	0	0	0
0.12	0.0002	0	0	0	0

$R_1^G(\overline{\Delta}_{1,n}, a),\ G \in \mathscr{C}_2(a)$

Values of a	$n = 100$	$n = 250$	$n = 500$	$n = 750$	$n = 1000$
0	0.8598	0.8648	0.8877	0.8688	0.86
0.03	0.9417	0.9692	0.9879	0.6488	0.083
0.05	0.9323	0.5986	0.0031	0	0
0.07	0.8064	0.0082	0	0	0
0.09	0.181	0	0	0	0
0.1	0.0867	0	0	0	0
0.12	0.0062	0	0	0	0

$R_1^G(\overline{\Delta}_{2,n}, a),\ G \in \mathscr{C}_1(a)$

Values of a	$n = 100$	$n = 250$	$n = 500$	$n = 750$	$n = 1000$
0	0.9228	0.8824	0.9699	0.9543	0.8085
0.03	0.9441	0.5673	0.067	0.0042	0.003
0.05	0.8593	0.2654	0	0	0
0.07	0.5849	0.0415	0	0	0
0.09	0.4318	0	0	0	0
0.1	0.3675	0	0	0	0
0.12	0.1286	0	0	0	0

$R_1^G(\overline{\Delta}_{2,n}, a),\ G \in \mathscr{C}_2(a)$

Values of a	$n = 100$	$n = 250$	$n = 500$	$n = 750$	$n = 1000$
0	0.9228	0.8824	0.9699	0.9543	0.8085
0.03	0.9641	0.0684	0	0	0
0.05	0.8987	0	0	0	0
0.07	0.6457	0	0	0	0
0.09	0.5214	0	0	0	0
0.1	0.2465	0	0	0	0
0.12	0.2294	0	0	0	0

First, note that for all cases, the error of the second kind decreases as n is increases. Second, if we fix a real number γ in (0, 1) which is an upper bound for the error of the second kind, the calculation of the error of the second kind in both tests gives the smallest value of a for which one can distinguish the alternatives from H_0; for example, for $n = 500$ and if we fix $\gamma = 0.1$, the corresponding a for the first test is 0.05 and that for the second test is 0.03. Furthermore, without fixing an upper bound γ and for large values of n, say $n \geqslant 500$, one must note that the errors of the second kind are always smaller when the distance **d** is based on the functional $S(G)$ than for the d_∞-distance: this would give the same conclusion as the theoretical results.

5. Proofs

5.1. Proof of Theorem 3.1

Let $\delta_n = (n/\log n)^{-1/(\gamma+N-1)}$.

5.1.1. Proof of (3.1)

Introduce a partition of $[0, 1]^{N-1}$ into M cubes Q_q, $q = 1, \ldots, M = (b_2^{1/\gamma}\delta_n)^{-(N-1)}$, with edges of length $b_2^{1/\gamma}\delta_n$. Assume without loss of generality that M is an integer. Let η be a function such that η is $\mathscr{C}^\infty$, $\eta(t) = 0$ if $t \notin [-\frac{1}{2}, \frac{1}{2}]^{N-1}$, $\eta(t) \geqslant 1$ if $t \in [-\frac{1}{2}, \frac{1}{2}]^{N-1}$, $\sup_t |\eta^{(k+1)}(t)| \leqslant L_1$ where $k = \lfloor \gamma \rfloor$, and write $\eta^* = \sup_{t \in [-\frac{1}{2},\frac{1}{2}]^{N-1}} \eta(t) \leqslant \infty$ and $\bar{\eta} = \int_{[-\frac{1}{2},\frac{1}{2}]^{N-1}} \eta(t)\mathrm{d}t$. For $q = 1, \ldots, M$, define the sets

$$G_0 = \{x = (x_N, y) \in [0, 1]^N : 0 \leqslant x_N \leqslant g_0(y)\},$$

$$G_q = \{x = (x_N, y) \in [0, 1]^N : 0 \leqslant x_N \leqslant g_q(y)\},$$

$$G_q^* = \{x = (x_N, y) \in [0, 1]^N : g_q(y) \leqslant x_N \leqslant g_0(y)\},$$

where $g_q(y) = g_0(y) - b_2\delta_n^\gamma \eta((y - u_q)/b_2^{1/\gamma}\delta_n)$. Denote by $\mathscr{G}_\infty^M$ the class of sets G_q for all $q = 1, \ldots, M$: it is clear that $\mathscr{G}_\infty^M$ is included in $\Lambda_{1,n}(b_2\delta_n^\gamma)$.

Set $\zeta_q^{(n)} = (\mathrm{d}P_q/\mathrm{d}P_0)(X_1, \ldots, X_n)$. Then, for any decision rule Δ_n,

$$P_{G_0}[\Delta_n = 1] + \sup_{G \in \Lambda_{1,n}(b_2\delta_n^\gamma)} P_G[\Delta_n = 0] \geqslant P_0[\Delta_n = 1] + \frac{1}{M}\sum_{q=1}^{M} P_q[\Delta_n = 0]$$

$$\geqslant (1 - \varepsilon)P_0\left[\frac{1}{M}\sum_{q=1}^{M} \zeta_q^{(n)} \geqslant (1 - \varepsilon)\right],$$

where P_G, P_0 and P_q respectively denote the probability distribution of the data when they are uniformly distributed on G, G_0 and G_q. The last inequality holds for any positive real $\varepsilon < 1$. Under H_0, and since $\operatorname{leb} G_q = \operatorname{leb} G_{q'}$ for all $q, q' = 1, \ldots, M$, we have

$$\zeta_q^{(n)} = \prod_{i=1}^{n} \frac{\text{leb}\, G_0\, I\{X_i \in G_q\}}{\text{leb}\, G_q\, I\{X_i \in G_0\}} = \left(\frac{\text{leb}\, G_0}{\text{leb}\, G_*}\right)^n \prod_{i=1}^{n} I\{X_i \in G_q\},$$

where $\text{leb}\, G_* = \text{leb}\, G_q$ for all $q = 1, \ldots, M$. Set $Z_M = (1/M)\sum_{q=1}^{M} \zeta_q^{(n)}$. If, for all $\varepsilon < 1$, $P_0[|Z_M - 1| \geqslant \varepsilon] \to 0$ as $n \to \infty$, then $P_0[Z_M \geqslant (1-\varepsilon)] \to 1$ as $n \to \infty$.

Note that $E_0[Z_M] = 1$ and that

$$\begin{aligned} E_0[Z_M^2] &= \frac{1}{M^2}\left(\frac{\text{leb}\, G_0}{\text{leb}\, G_*}\right)^{2n} E_0\left[\left(\sum_{q=1}^{M} I\{X_i \in G_q, \forall i = 1, \ldots, n\}\right)^2\right] \\ &= \frac{1}{M^2}\left(\frac{\text{leb}\, G_0}{\text{leb}\, G_*}\right)^{2n}\left(E_0\left[\sum_{q=1}^{M} I\{X_i \in G_q, \forall i = 1, \ldots, n\}\right]\right. \\ &\qquad \left. + E_0\left[\sum_{q \neq q'} I\{X_i \in G_q \cap G'_q, \forall i = 1, \ldots, n\}\right]\right). \end{aligned}$$

Set

$$T_1 = \frac{1}{M^2}\left(\frac{\text{leb}\, G_0}{\text{leb}\, G_*}\right)^{2n} E_0\left[\sum_{q=1}^{M} \mathrm{I}\{X_i \in G_q, \forall i = 1, \ldots, n\}\right]$$

and note that

$$T_1 = \frac{1}{M}\left(\frac{\text{leb}\, G_0}{\text{leb}\, G_*}\right)^{n}.$$

Also set

$$T_2 = \frac{1}{M^2}\left(\frac{\text{leb}\, G_0}{\text{leb}\, G_*}\right)^{2n} E_0\left[\sum_{q \neq q'} I\{X_i \in G_q \cap G'_q, \forall i = 1, \ldots, n\}\right]$$

and note that

$$\begin{aligned} T_2 &= \frac{1}{M^2}\left(\frac{\text{leb}\, G_0}{\text{leb}\, G_*}\right)^{2n} \sum_{q \neq q'} (E_0[I\{X_1 \in G_q \cap G_{q'}\}])^n \\ &= \frac{1}{M^2}\left(\frac{\text{leb}\, G_0}{\text{leb}\, G_*}\right)^{2n} \sum_{q \neq q'} \left(\frac{\text{leb}\, G_0 - \text{leb}\, G^*_{q'}\, \text{leb}\, G^*_q}{\text{leb}\, G_0}\right)^n, \end{aligned} \tag{5.1}$$

where (5.1) is due to the independence of the variables X_i. Write $\text{leb}\, G^* = \text{leb}\, G^*_q$ for any $q \in \{1, \ldots, M\}$ since $\text{leb}\, G^*_q = \text{leb}\, G^*_p$ for all $p, q \in \{1, \ldots, M\}$. Then

$$E_0[Z_M^2] = \frac{1}{M}\left(\frac{\text{leb } G_0}{\text{leb } G_*}\right)^n + \frac{1}{M^2}\left(\frac{\text{leb } G_0}{\text{leb } G_*}\right)^{2n} \sum_{q \neq q'} \left(\frac{\text{leb } G_0 - \text{leb } G^*_{q'} - \text{leb } G^*_q}{\text{leb } G_0}\right)^n. \qquad (5.2)$$

Write $\bar{g}_0 = \text{leb } G_0$; for b_2 small enough, note that

$$\left(\frac{\text{leb } G_0}{\text{leb } G_*}\right)^n = (1 - b_2^{1+(N-1)/\gamma}\delta_n^{\gamma+N-1}(\bar{\eta}/\bar{g}_0))^{-n} = n^{(b_2^{1+(N-1)/\gamma}(\bar{\eta}/\bar{g}_0))}(1 + o(1)).$$

If b_2 is small enough and satisfies as $n \to \infty$, $b_2^{1+(N-1)/\gamma}(\bar{\eta}/\bar{g}_0) < (N - 1/\gamma + N - 1)$, we obtain

$$\frac{1}{M}\left(\frac{\text{leb } G_0}{\text{leb } G_*}\right)^n \to 0. \qquad (5.3)$$

In the same way, for b_2 small enough, we obtain the following approximation for the second part of (5.2):

$$\begin{aligned}
&\frac{1}{M^2}\left(\frac{\text{leb } G_0}{\text{leb } G_*}\right)^{2n} \sum_{q \neq q'} \left(\frac{\text{leb } G_0 - \text{leb } G^*_{q'} - \text{leb } G^*_q}{\text{leb } G_0}\right)^n \\
&\quad = \frac{M^2 - M}{M^2}(1 - b_2^{1+(N-1)/\gamma}\delta_n^{\gamma+N-1}(\bar{\eta}/\bar{g}_0))^{-2n}(1 - 2b_2^{1+(N-1)/\gamma}\delta_n^{\gamma+N-1}(\bar{\eta}/\bar{g}_0))^n \\
&\quad = \frac{M^2 - M}{M^2}(1 + o(1)) \to 1, \qquad n \to \infty.
\end{aligned} \qquad (5.4)$$

From (5.2), (5.3), (5.4) and by Chebyshev's inequality, (3.1) holds.

5.1.2. Proof of (3.2)

First, consider the error of the second kind. For all G in $\Lambda_{1,n}(b_3\delta_n^\gamma)$,

$$\begin{aligned}
P_G[\overline{\Delta}_{1,n} = 0] &\leqslant P_G[d_\infty(G, G_0) - d_\infty(G_n, G) < C_1\delta_n^\gamma] \\
&\leqslant P_G[d_\infty(G_n, G) > (b_3 - C_1)\delta_n^\gamma].
\end{aligned} \qquad (5.5)$$

As soon as there exists a constant b_3 such that $b_3 - C_1$ is large enough, and using relation (2.15) in Korostelev and Tsybakov (1993b), we obtain

$$P_G[\overline{\Delta}_{1,n} = 0] \leqslant n^{-p(b_3 - C_1)}, \qquad (5.6)$$

where p is an arbitrary fixed positive number. It follows that $R_1(\overline{\Delta}_{1,n}, b_3\delta_n^\gamma)$ is asymptotically equal to zero.

Under P_{G_0}, providing some upper bound for $R_0(\overline{\Delta}_{1,n})$ reduces to provide some bound for $P_G[d_\infty(G_n, G) \geqslant C_1\delta_n^\gamma]$, for all $G \in \mathscr{G}$. Relation (3.2) follows since the last inequality is similar to the right-hand side of relation (5.5).

5.2. Proof of Theorem 3.2

Let $\psi_n = n^{-[\gamma+(N-1)/2]/(\gamma+N-1)}$ and $\delta_n = n^{-1/(\gamma+N-1)}$.

5.2.1. Proof of (3.3)

Consider the partition defined in the proof of (3.1) with b_4 in place of b_2 and $\delta_n = n^{-(1/\gamma+N-1)}$, and consider also the function η defined in the proof of Theorem 3.1. Let $\omega = (\omega_1, \ldots, \omega_M)$ be a binary vector such that $\omega_1, \ldots, \omega_M$ are independent and identically distributed Bernoulli random variables and set $\Omega = \{\omega : \sum_{q=1}^{M} w_q = M^{1/2}\}$; assume without loss of generality that $M^{1/2}$ is an integer. Write $|\Omega| = \operatorname{card}\Omega$ and define the sets

$$G_0 = \{x = (x_N,\ y) \in [0,\ 1]^N : 0 \leqslant x_N \leqslant g_0(y)\},$$
$$G_\omega = \{x = (x_N,\ y) \in [0,\ 1]^N : 0 \leqslant x_N \leqslant g_\omega(y)\},$$

where

$$g_\omega(y) = g_0(y) - b_4\delta_n^\gamma \sum_{\omega\in\Omega} \omega_q \eta\left(\frac{y-u_q}{b_4^{1/\gamma}\delta_n}\right).$$

Set $\mathscr{G}^M = \{G_\omega : \omega \in \Omega\}$, the parametric class of G_ω. It is clear that $\mathscr{G}^M$ is included in $\Lambda_{2,n}(b_4\psi_n)$, with $\psi_n = n^{-[\gamma+(N-1)/2]/(\gamma+N-1)}$.

Set $\bar{P} = (1/|\Omega|)\sum_{\omega\in\Omega} P_{G_\omega}$. Thus, for any decision rule Δ_n, we obtain

$$\begin{aligned} P_{G_0}[\Delta_n = 1] &+ \sup_{G\in\Lambda_{2,n}(b_4\psi_n)} P_G[\Delta_n = 0] \\ &\geqslant P_0[\Delta_n = 1] + \bar{P}[\Delta_n = 0] \\ &\geqslant \frac{1}{|\Omega|}\sum_{\omega\in\Omega} P_\omega\left[\frac{1}{|\Omega|}\sum_{\omega'\in\Omega}\frac{\mathrm{d}P_{\omega'}}{\mathrm{d}P_0} \leqslant \frac{1}{\varepsilon}\right], \end{aligned} \tag{5.7}$$

where P_ω and P_0 denote the probability distribution of the data when their underlying support is G_ω, $\omega \in \Omega$, and G_0 respectively, and ε is an arbitrary positive real. For simplicity's sake, denote $\bar{g}_0 = \operatorname{leb}(G_0)$. We first fix $\omega \in \Omega$, and by Chebyshev's inequality we obtain

$$\begin{aligned} P_\omega\left[\frac{1}{|\Omega|}\sum_{\omega'\in\Omega}\frac{\mathrm{d}P_{\omega'}}{\mathrm{d}P_0} \leqslant \frac{1}{\varepsilon}\right] &\geqslant 1 - \frac{\varepsilon}{|\Omega|}\sum_{\omega'\in\Omega} E_\omega\left[\frac{\mathrm{d}P_{\omega'}}{\mathrm{d}P_0}\right] \\ &= 1 - \frac{\varepsilon}{|\Omega|}\sum_{\omega'\neq\omega:\omega,\omega'\in\Omega} E_\omega\left[\frac{\mathrm{d}P_{\omega'}}{\mathrm{d}P_0}\right] - \frac{\varepsilon}{|\Omega|}E_\omega\left[\frac{\mathrm{d}P_\omega}{\mathrm{d}P_0}\right]. \end{aligned} \tag{5.8}$$

Since b_4 is chosen small enough, the final term on the right in (5.8) becomes

$$E_\omega\left[\frac{\mathrm{d}P_\omega}{\mathrm{d}P_0}\right] = \left(\frac{\operatorname{leb} G_0}{\operatorname{leb} G_\omega}\right)^n = \exp(M^{1/2}b_4(\bar{\eta}/\bar{g}_0))(1+o(1)). \tag{5.9}$$

Furthermore, $|\Omega|$ is equal to the number of ways of choosing $M^{1/2}$ elements from a set of M elements, that is, $|\Omega| = C^M_{M^{1/2}}$. Note that there exist two positive constants $\tilde{c}$ and c' such that

$$\tilde{c} n^{n+(1/2)} \exp(-n) \leqslant n! \leqslant c' n^{n+(1/2)} \exp(-n) \qquad \forall n. \tag{5.10}$$

Then, as n goes to infinity,

$$\begin{aligned} C^M_{M^{1/2}} &= \exp[M \log M - \tfrac{1}{2} M^{1/2} \log M - (M - M^{1/2}) \log(M - M^{1/2})](1 + o(1)) \\ &= \exp[\tfrac{1}{2} M^{1/2} \log M](1 + o(1)). \end{aligned}$$

This entails

$$\frac{\varepsilon}{|\Omega|} E_\omega \left[\frac{\mathrm{d}P_\omega}{\mathrm{d}P_0} \right] \to 0, \qquad \text{as } n \to \infty. \tag{5.11}$$

Now consider the case $\omega' \neq \omega$, ω, $\omega' \in \Omega$. Since $\operatorname{leb} G_\omega = \operatorname{leb} G_{\omega'}$, for all ω, $\omega' \in \Omega$,

$$E_\omega \left[\frac{\mathrm{d}P_{\omega'}}{\mathrm{d}P_0} \right] = \left(\frac{\operatorname{leb}(G_{\omega'} \cap G_\omega) \operatorname{leb} G_0}{(\operatorname{leb} G_\omega)^2} \right)^n. \tag{5.12}$$

Define

$$l = \sharp\{ q = 1, \ldots, M : g_\omega(y) = g_{\omega'}(y) \neq g_0(y) \forall y \in Q_q \}.$$

Since ω and ω' belong to Ω, l cannot exceed $M^{1/2} - 1$. Note that $\operatorname{leb}(G_{\omega'} \cap G_\omega) = g_0(y) - 2M^{1/2}(b_4 \bar{\eta}/n) + l(b_4 \bar{\eta}/n)$. Then, as n goes to infinity, (5.12) can be written as

$$\begin{aligned} \left(\frac{\operatorname{leb}(G_{\omega'} \cap G_\omega) \operatorname{leb} G_0}{(\operatorname{leb} G_\omega)^2} \right)^n &= \left(\frac{1 - 2M^{1/2}(b'_4/n) + l(b'_4/n)}{(1 - M^{1/2} b'_4/n)^2} \right)^n \\ &= (1 + (l/n) b'_4)^n (1 + o(1)) \\ &= \exp(l b'_4)(1 + o(1)), \end{aligned} \tag{5.13}$$

where $b'_4 = (b_4 \bar{\eta})/\bar{g}_0$ is a positive constant. From (5.12) and (5.13), we obtain

$$\sum_{\omega' \neq \omega : \omega, \omega' \in \Omega} E_\omega \left[\frac{\mathrm{d}P_{\omega'}}{\mathrm{d}P_0} \right] = \sum_{l=1}^{M^{1/2}-1} \sum_{\omega' \in \Omega_l} \exp(l b'_4)(1 + o(1)),$$

where $\Omega_l = \{ \omega' \in \Omega : \operatorname{leb}(G_\omega \cap G_{\omega'}) = l \}$. Note that $\operatorname{card} \Omega_l = C_l^{M^{1/2}} C^{M - M^{1/2}}_{M^{1/2} - l}$. For n large enough and from (5.10), we have

$$C_l^{M^{1/2}} C_{M^{1/2}-l}^{M-M^{1/2}} = \exp[\tfrac{1}{2}M^{1/2}\log M - l\log l - 2(M^{1/2}-l)\log(M^{1/2}-l)$$
$$+ (M - M^{1/2})\log(M - M^{1/2})$$
$$- (M - 2M^{1/2} + l)\log(M - 2M^{1/2} + l)](1 + o(1))$$
$$\leqslant \exp[\tfrac{1}{2}M^{1/2}\log M - l\log l](1 + o(1)). \tag{5.14}$$

Thus, for n large enough

$$\frac{\varepsilon}{|\Omega|}\sum_{\omega'\neq\omega:\omega,\omega'\in\Omega} E_\omega\left[\frac{\mathrm{d}P_{\omega'}}{\mathrm{d}P_0}\right] = \varepsilon\sum_{l=1}^{M^{1/2}-1}\exp(-l\log l)(1+o(1)) < \infty. \tag{5.15}$$

There exists ε^* such that for all $\varepsilon < \varepsilon^*$, $1 - \varepsilon\sum_{l=1}^{\infty}\exp(-l\log l)) > 0$. Set $\beta^* = \varepsilon^*(1 - \sum_{l=1}^{\infty}\exp(-l\log l)\varepsilon^*)$. From (5.8), (5.11), (5.15) and for all $\beta < \beta^*$, (3.3) holds.

5.2.2. Proof of (3.4)

Choose $\beta_1 > 0$ and $\beta_2 > 0$ such that $\beta = \beta_1 + \beta_2$. First consider the risk of the second kind. Choose $b_5 > C_2$; $\forall G \in \Lambda_{2,n}(b_5\psi_n)$, we have

$$P_G[\overline{\Delta}_{2,n} = 0] \leqslant P_G[(S_n - S(G))^2 > (b_5 - C_2)^2 n^{-(2\gamma+N-1)/(\gamma+N-1)}]. \tag{5.16}$$

By Chebyshev's inequality, we obtain

$$P_G[\overline{\Delta}_{2,n} = 0] \leqslant \frac{E_G[(S_n - S(G))^2]}{(b_5 - C_2)^2 n^{-(2\gamma+N-1)/(\gamma+N-1)}}. \tag{5.17}$$

For n large enough, adapting the results on support estimation in Gayraud (1997), there exists a constant $b_5 > C_2$ such that $R_1(\overline{\Delta}_{2,n}, b_5 n^{-(2\gamma+N-1)/[2(\gamma+N-1)]})$ is bounded from above by β_1.

Under P_{G_0}, the proof of $R_0(\overline{\Delta}_{2,n}) \leqslant \beta_2$ is reduced to proving that $P_G[(S_n - S(G))^2 \geqslant C_2 n^{-(2\gamma+N-1)/(\gamma+N-1)}] \leqslant \beta_2$, $\forall G \in \mathscr{G}$. Noting that the last inequality is similar to inequality (5.16), relation (3.4) is then satisfied.

5.3. Proof of Theorem 3.3

Let $\delta_n = n^{-1/(\gamma+N-1)}$ and $\psi_n = n^{-[\gamma+(N-1)/2]/(\gamma+N-1)}$.

5.3.1. Proof of (3.5)

As in the proof of Theorem 3.2, consider the set $\omega = \{\omega : \sum_{q=1}^{M} w_q = M^{1/2}\}$, where M is the number of cubes of the partition of $[0, 1]^{N-1}$, and

$$G_0 = \{x = (x_N, y) \in [0,1]^N : 0 \leqslant x_N \leqslant g_0(y)\},$$
$$G_\omega = \{x = (x_N, y) \in [0,1]^N : 0 \leqslant x_N \leqslant g_\omega(y)\},$$

where

$$g_\omega(y) = g_0(y) - b_6\delta_n^\gamma \sum_{\omega\in\Omega} \omega_q \eta\left(\frac{y-u_q}{b_6^{1/\gamma}\delta_n}\right).$$

Note that the parametric family of sets $\{G_\omega : \omega \in \Omega\}$ is included in $\Lambda_{3,n}(b_6\psi_n)$. Thus, (3.5) follows from the proof of (3.3).

5.3.2. Proof of (3.6)

Choose $\beta_1 > 0$ and $\beta_2 > 0$ such that $\beta = \beta_1 + \beta_2$. Consider first the error of the second kind. Choose $b_7 > C_3$; then for all $G \in \Lambda_{3,n}(b_7\psi_n)$,

$$P_G[\overline{\Delta}_{3,n} = 0] \leqslant P_G[|T(G) - T_n| > (b_7 - C_3)\psi_n],$$
$$\leqslant E_G[(T(G) - T_n)^2 \psi_n^{-2}(b_7 - C_3)^{-2}].$$

Since $b_7 > C_3$ and from Theorem 2 in Gayraud (1997), the relation $E_G[(T(G) - T_n)^2 \times \psi_n^{-2}(b_7 - C_3)^{-2}] \leqslant \beta_1$, as $n \to \infty$ is satisfied.

Under P_{G_0}, G is G_0 and then $T(G_0) = \int_{[0,1]^{N-1}} |g_0(y) - g_0(y)|\mathrm{d}y$ is equal to zero. Then to prove that $R_0(\overline{\Delta}_{3,1}) \leqslant \beta_2$, as $n \to \infty$, reduces to proving that $P_G[|T_n - T(G)| \geqslant C_3 n^{-[\gamma+(N-1)/2]/(\gamma+N-1)}] \leqslant \beta_2$ as $n \to \infty$, for $G \in \mathscr{G}$. Then (3.6) holds.

Acknowledgments

I thank Alexander B. Tsybakov for useful discussion and helpful suggestions.

References

Baraud, Y., Huet, S. and Laurent, B. (1999) Adaptive tests of linear hypotheses by model selection. Preprint, École Normale Supérieure, Paris.

Birgé, L. (1983) Approximation dans les espaces métriques et théorie de l'estimation. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **65**, 181–237.

Burnashev, M.V. (1979) On the minimax detection of an inaccurately known signal in a white noise background. *Theory Probab. Appl.*, **24**, 107–119.

Donoho, D.L., Johnstone, I.M., Kerkyacharian, G. and Picard, D. (1995) Wavelet shrinkage: Asymptopia? *J. Roy. Statist. Soc. Ser. B*, **57**, 301–369.

Ermakov, M.S. (1990a) Minimax detection of a signal in a white Gaussian noise. *Theory Probab. Appl.*, **35**, 667–679.

Ermakov, M.S. (1990b) A minimax test for hypotheses on a spectral density. *Zap. Nauchn. Sem. Leningr. Otdel. Mat. Inst. Steklov.*, **184**, 115–125.

Ermakov, M.S. (1990c) Asymptotic minimaxity of usual goodness of fit tests. In B. Grigelionis, Yu. V. Prohorov, V.V. Sazonov and V. Statulevicius (eds), *Probability Theory and Mathematical Statistics: Proceedings of the Fifth Vilnius Conference*, Vol. 1, pp. 323–331. Utrecht: VSP.

Ermakov, M.S. (1994) Minimax nonparametric testing of hypotheses on the distribution density. *Theory Probab. Appl.*, **39**, 396–416.

Ermakov, M.S. (1996) Asymptotically minimax criteria for testing composite nonparametric hypotheses. *Problems Inform. Transmission*, **32**, 184–196.

Feldmann, D., Härdle, W., Hafner, C., Hoffmann, M., Lepski, O. and Tsybakov, A.B. (1998) Testing linearity in a stochastic volatility model. Preprint SFB 373, Humboldt Universität zu Berlin.

Gayraud, G. (1997) Estimation of functionals of density support. *Math. Methods Statist.*, **6**, 26–46.

Gayraud, G. and Tsybakov, A.B. (1999) Testing hypotheses about contours in images. Research Report no. 512, Laboratoire de Probabilités et Modèles Aléatoires, Universités de Paris 6 et Paris 7.

Härdle, W. and Kneip, A. (1999) Testing a regression model when we have smooth alternatives in mind. *Scand. J. Statist.*, **26**, 221–238.

Härdle, W., Sperlich, S. and Spokoiny, V.G. (1997) Semiparametric single index versus fixed link function modelling. *Ann. Statist.*, **25**, 212–243.

Ibragimov, I.A. and Has'minskii, R.Z. (1977) One problem of statistical estimation in Gaussian white noise. *Soviet Math. Dokl.*, **236**, 1351–1354.

Ingster, Yu.I. (1982) Minimax nonparametric detection of a signal in Gaussian white noise. *Problems Inform. Transmission*, **18**, 130–140.

Ingster, Yu.I. (1984) Asymptotical minimax testing hypotheses on the distribution density of an independent sample. *Zap. Nauchn. Sem. Leningr. Otdel. Mat. Inst. Steklov.*, **136**, 74–96.

Ingster, Yu.I. (1986a) Minimax testing of nonparametric hypotheses on a distribution density in L_p-metrics. *Theory Probab. Appl.*, **31**, 333–337.

Ingster, Yu.I. (1986b) Asymptotically minimax testing of nonparametric hypotheses. In Yu.V. Prohorov, V.A. Statulevicius, V.V. Sazonov and B. Grigelionis (eds), *Probability Theory and Mathematical Statistics, Proceedings of the Fourth Vilnius Conference*, Vol. 1, pp. 553–574. Utrecht: VSP.

Ingster, Yu.I. (1990) Minimax detection of signals in L_p-metrics. *Zap. Nauchn. Sem. Leningr. Otdel. Mat. Inst. Steklov.*, **184**, 152–168.

Ingster, Yu.I. (1993a) Asymptotically minimax hypothesis testing for nonparametric alternatives, I. *Math. Methods Statist.*, **2**, 85–114.

Ingster, Yu.I. (1993b) Asymptotically minimax hypothesis testing for nonparametric alternatives, II. *Math. Methods Statist.*, **3**, 171–189.

Ingster, Yu.I. (1993c) Asymptotically minimax hypothesis testing for nonparametric alternatives, III. *Math. Methods Statist.*, **4**, 249–268.

Ingster, Yu.I. (1994) Minimax testing of hypotheses on the distribution density for ellipsoids in l_p. *Theory Probab. Appl.*, **39**, 417–436.

Korostelev, A.P. and Tsybakov, A.B. (1993a) *Minimax Theory of Image Reconstruction*, Lecture Notes in Statist. 82. New York: Springer-Verlag.

Korostelev, A.P. and Tsybakov, A.B. (1993b) Estimation of the density support and its functionals. *Problems Inform. Transmission*, **29**, 3–18.

Lepski, O.V. (1993) On asymptotically exact testing of nonparametric hypotheses. Discussion paper, Université de Louvain.

Lepski, O.V. and Spokoiny, V.G. (1999) Minimax nonparametric hypothesis testing: the case of an inhomogeneous alternative. *Bernoulli*, **5**, 333–358.

Lepski, O.V. and Tsybakov, A.B. (1996) Asymptotically exact nonparametric hypothesis testing in sup-norm and at a fixed point. Preprint SBF 373, Humboldt-Universität zu Berlin.

Pouet, C. (1999) On testing nonparametric hypotheses for analytic regression functions in Gaussian noise. *Math. Methods Statist.*, **8**, 536–549.

Pouet, C. (2000) Tests minimax non-paramétriques: hypothèse nulle composite et constantes exactes. Doctoral thesis, Université Paris VI.

Spokoiny, V.G. (1996) Adaptive hypothesis testing using wavelets, *Ann. Statist.*, **24**, 2477–2498.

Spokoiny, V.G. (1997) Testing a linear hypothesis using Haar transform. Research Report no. 314, WIAS, Berlin.

Spokoiny, V.G. (1998) Adaptive and spatial testing of a nonparametric hypothesis. *Math. Methods Statist.*, **7**, 245–273.

Suslina, I.A. (1993) Minimax detection of a signal for l_p-ball. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.*, **207**, 127–137.

Received February 1997 and revised November 2000