Weierstraß–Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

On distinguishability of two nonparametric sets of hypothesis

Michael S. Ermakov

submitted: 10th October 1996

Mechanical Engineering
Problems Institute
Russian Academy of Sciences
Bolshoy pr., V.O., 61
199178 St. Petersburg
Russia

Preprint No. 273
Berlin 1996

1991 Mathematics Subject Classification. 62G10; Secondary 62G20.
Key words and phrases. Hypothesis testing, asymptotic efficiency, signal detection, hypothesis testing about density, nonparametric hypothesis testing.

Edited by
Weierstraß–Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
D — 10117 Berlin
Germany

Fax: + 49 30 2044975
e–mail (X.400): c=de;a=d400-gw;p=WIAS-BERLIN;s=preprint
e–mail (Internet): preprint@wias-berlin.de

Abstract. Let we observe a signal $S(t)$, $t \in (0,1)$ in Gaussian white noise $\epsilon\, dw(t)$. The problem is to test a hypothesis $S \in \Theta_1 \subset L_2(0,1)$ versus alternatives $S \in \Theta_2 \subset L_2(0,1)$. The sets Θ_1, Θ_2 are closed and bounded. We show that there exists a statistical procedure allowing to make a true solution $S \in \Theta_1$ or $S \in \Theta_2$ with probability tending to one as $\epsilon \to 0$ (i.e. to distinguish two nonparametric sets Θ_1 and Θ_2) iff there exists a finite-dimensional subspace $H \subset L_2(0,1)$ such that the projections Θ_1 and Θ_2 on H have no common points. A similar result is also obtained for the problems of testing hypotheses about density.

1. General Setting and Main Result. In estimation problems an asymptotic behaviour of estimators is usually characterised from the three different viewpoints: consistency, rate of convergence and efficiency. In hypothesis testing the analogous of these viewpoints are the distinguishability of the hypotheses and the alternatives, the optimal rates of distinguishability (the optimal rates of approaching the hypotheses and the alternatives allowing to distinguish them) and the asymptotic minimax or Bayes optimality of tests.

For practical applications with a finite number of parameters the problems of distinguishability and optimal rates of distinguishability do not present any difficulty. Essential difficulties arise if the sets of hypotheses or alternatives have nonparametric nature. For example, such a problem is testing a hypothesis that L_2-norm of deviation of density from the density of uniform distribution exceeds $\rho > 0$ or a similar problem of signal detection in Gaussian white noise. Here testing a hypothesis is possible only if additional a priori information is available. Thus the problem of distinguishability in nonparametric setting deserves a special investigation. For the first time testing nonparametric hypotheses has been considered by Mann and Wald (1942) and Stein (1956). In these papers the basic settings have been proposed. The last years the problem of testing nonparametric hypotheses was investigating intensively (see Ermakov (1990),(1995), Ingster (1988),(1993) and references therein). The main attention has been paid to the optimal rates of distinguishability and to construction of asymptotically minimax sequences of tests. At the same time the simplest distinguishability problem has been considered only as an auxiliary question or as in the context of more powerful results on optimal rates of distinguishability.

First of all, among the results on distinguishability problem, we should mention Burnashev's paper (1979). For a signal observed in the Gaussian white noise Burnashev has shown that the problem of nonparametric signal detection in L_2-norm cannot be solved without additional a priori information. A similar result for the problem of testing hypotheses about density has been obtained by Ingster (1993). For a signal detection in the Gaussian noise, the problem of distinguishability has been studied in Ermakov (1990) for the sets of alternatives represented as difference of two ellipsoids in L_2. A similar setting for a difference of two l_p bodies has been considered in Ingster (1993).

The purpose of the paper is to find necessary and sufficient conditions of distinguishability of two bounded sets of hypotheses for the two problems: signal detection in the Gaussian white noise and testing hypotheses about density. As we know, almost all widespread statistical models are usually reduced to their analogies with Gaussian white noise (see Donoho and Liu (1987), Brown and Low (1992), Nussbaum (1995)). By this reason, we first of all consider the model of signal detection in the Gaussian white noise. Then a similar model of testing hypotheses about density will be considered.

Suppose we observe a random process $Y(t)$, $t \in (0,1)$, defined by the stochastic differential equation

$$dY(t) = S(t)dt + \epsilon\, dw(t), \qquad \epsilon > 0.$$

Here $S(t)$ is an unknown signal and $dw(t)$ is the Gaussian white noise. The problem is to test the hypothesis $S \in \Theta_1 \subset L_2(0,1)$ versus the alternative $S \in \Theta_2 \subset L_2(0,1)$. The sets Θ_1 and Θ_2 are assumed to be closed and bounded.

For a test K_ϵ, denote by $\alpha_{\theta,\epsilon}(K_\epsilon)$ its type I error probability for the hypothesis $\theta \in \Theta_1$, and by $\beta_{\theta,\epsilon}(K_\epsilon)$ its type II error probability for the alternative $\theta \in \Theta_2$. Let

$$\alpha_\epsilon(K_\epsilon) = \sup_{\theta\in\Theta_1} \alpha_{\theta\epsilon}(K_\epsilon), \quad \beta_\epsilon(K_\epsilon) = \sup_{\theta\in\Theta_2} \beta_{\theta\epsilon}(K_\epsilon)$$

We say that the sets of hypotheses Θ_1 and the sets of alternatives Θ_2 are distinguishable if there exists a family of tests K_ϵ such that

$$\limsup_{\epsilon\to 0}\, (\alpha_\epsilon(K_\epsilon) + \beta_\epsilon(K_\epsilon)) < 1.$$

Otherwise we shall say that the sets of hypotheses and alternatives are indistinguishable.

The problem of distinguishability admits the following interpretation if the sets of hypotheses and alternatives converge to each other. Assume we have two families of sets $\Theta_1(\rho)$ and $\Theta_2(\rho)$ with $\rho \in R^1_+$ such that $\Theta_1(\rho_2) \subset \Theta_1(\rho_1)$, $\Theta_2(\rho_2) \subset \Theta_2(\rho_1)$ for all $0 < \rho_1 < \rho_2 < \infty$. Let $\Theta_1(\rho) \cap \Theta_2(\rho) = \emptyset$ for all $\rho > 0$ and let $\Theta_1(0) \cap \Theta_2(0) \neq \emptyset$. We call ρ_ϵ the optimal rate of distinguishability if the sets $\Theta_1(\rho_\epsilon)$ and $\Theta_2(\rho_\epsilon)$ are distinguishable and for any $\rho_{1\epsilon}$, $\rho_{1\epsilon}/\rho_\epsilon \to 0$ as $\epsilon \to 0$, the sets $\Theta_1(\rho_{1\epsilon})$, $\Theta_2(\rho_{1\epsilon})$ are indistinguishable. Finally, we shall say that the families of sets $\Theta_1(\rho)$ and $\Theta_2(\rho)$ are asymptotically distinguishable if there exists $\rho_\epsilon \to 0$ as $\epsilon \to 0$ such that the sets $\Theta_1(\rho_\epsilon)$ and $\Theta_2(\rho_\epsilon)$ are distinguishable. It is clear that the families of sets $\Theta_1(\rho)$ and $\Theta_2(\rho)$ are asymptotically distinguishable if the sets $\Theta_1(\rho)$, $\Theta_2(\rho)$ are distinguishable for any fixed $\rho > 0$. Thus the problem of asymptotic distinguishability reduces easily to the simpler problem of the usual distinguishability.

The interest to the problem of asymptotic distinguishability can be illustrated by the following simple example of sets of hypotheses and alternatives. Let $\Theta_1(\rho) = \{0\}$ for all $\rho > 0$ and let $\Theta_2(\rho) = U \backslash B_\rho(0)$ where U is a closed bounded set and $B_\rho(0)$ is the ball in L_2 of the radius ρ centered at zero. As follows from Theorem 1.1 below the hypothesis $\theta = 0$ and the set of alternatives $\Theta_2(\rho)$ are distinguishable iff U is a compact set. This example was the starting point in the consideration of the problem.

We shall consider the problem of testing nonparametric hypothesis about density in the following setting. Let $X_1, \dots, X_n$ be i.i.d.r.v.'s on a probability space $(\Omega, \Im, P)$ and let the measure P be absolutely continuous w.r.t. a probability measure ν with the density $S(x) = dP/d\nu(x)$. The problem is to test the hypothesis $S \in \Theta_1 \subset L_2(\nu)$ versus the alternative $S \in \Theta_2 \subset L_2(\nu)$. For such a setting we can preserve the same notations and definitions as in the problem of signal detection. The only difference is that the parameter ϵ in the notation should now be replaced by the parameter n and instead of the asymptotics $\epsilon \to 0$ we consider the asymptotics $n \to \infty$.

We shall use the following notation. Let H be a subspace of $L_2(\nu)$. Denote by Π_H the projection operator on the subspace H and by $\dim(H)$ the dimension of H if H is finite-dimensional. For any $S_1,S_2 \in L_2(\nu)$ define the inner product

$$(S_1, S_2) = \int_\Omega S_1(t)S_2(t)\,\nu(dt)$$

and let $||S_1||^2 = (S_1, S_1)$. For any pair of subspaces $H_1, H_2 \subset L_2(\nu)$ denote $H_1 + H_2 = \{S : S = S_1 + S_2,\ S_1 \in H_1,\ S_2 \in H_2\}$.

The results for both the models are the same and are given below in Theorem 1.1. In this theorem, in the case of signal detection ν stands for the Lebesgue measure in $\Omega = (0,1)$.

Theorem 1.1. *Let Θ_1 and Θ_2 be closed bounded sets in $L_2(\nu)$. Then, both in the problem of signal detection and testing hypotheses about density, the sets Θ_1 and Θ_2 are distinguishable iff there exists a finite-dimensional subspace $H \subset L_2(\nu)$ such that $\Pi_H\Theta_1 \cap \Pi_H\Theta_2 = \emptyset$.*

Theorem 1.1 implies that each solvable problem of hypotheses testing has a "parametric counterpart".

Remark. For the problems of testing nonparametric hypotheses Theorem 1.1 can be considered as the analogy to the following result on nonparametric estimation. Let Θ be a bounded set in $L_2(\nu)$. Then there exists a consistent estimator of signals $S \in \Theta$ iff the closure of Θ is a compact set in $L_2(\nu)$ (see Ibragimov and Khasminskii (1977)).

First we shall prove Theorem 1.1 for the Gaussian white noise model. Then we shall point out the modifications in the proof required by the case of testing hypotheses about density.

2. Proof of Theorem 1.1. Signal in the Gaussian white noise. The sufficiency of condition is clear since under this condition the problem reduces to its finite-dimensional version.

The necessity will be proved separately for the following three cases: $\Theta_1 = \{S_0\}$, Θ_1 is a compact set, and Θ_1 is an arbitrary bounded set. The first two cases are considered to make transparent the idea of the proof for the most general case.

Lemma 2.1. *Let $\Theta_1 = \{S_0\}$ and let Θ_2 be a closed bounded subset of $L_2(\nu)$. Assume the hypothesis $S = S_0$ and the alternative $S \in \Theta_2$ are distinguishable. Then there exists a finite-dimensional subspace $H \in L_2(\nu)$ such that $\Pi_H S_0 \notin \Pi_H\Theta_2$.*

Clearly we can take the aforementioned H in such a way that $S_0 \in H$.

Proof of Lemma 2.1. Suppose the opposite. Then, for any sequence $\rho_k > 0$, $\rho_k \to 0$ as $k \to \infty$, there exists a sequence $S_1, S_2, \ldots \in \Theta_2$ such that $\sup\{|(S_j - S_0, S_k - S_0)| : 0 \le j < k\} < \rho_k$.

Fix a sequence $\epsilon_m > 0$ and define a sequence of a priori Bayes measures μ_m in such a way that $\mu_m(S_j) = 1/m$, $1 \leq j \leq m$. The corresponding Bayes likelihood ratios are

$$I_m = \frac{1}{m}\sum_{j=1}^{m} \exp\{\epsilon_m^{-1}\int_0^1 (S_j(t) - S_0(t))\,dw(t) - \frac{1}{2}\epsilon_m^{-2}||S_j - S_0||^2\} \qquad (2.1)$$

To complete the proof, it suffices to show that there exists a sequence $\rho_m \to 0$ as $m \to \infty$ such that

$$\lim_{m\to\infty} Var_{\mu_m}[I_m] = 0 \qquad (2.2)$$

By direct calculations we get

$$Var_{\mu_m}[I_m] = m^{-2} \sum_{j_1,j_2=1}^{m} \exp\{\epsilon_m^{-2}(S_{j_1} - S_0, S_{j_2} - S_0)\} - 1 \qquad (2.3)$$

which implies (2.2).

Lemma 2.2. *Let Θ_1 be a compact set in $L_2(\nu)$ and Θ_2 be a closed bounded subset of $L_2(\nu)$. Then Θ_1 and Θ_2 are distinguishable only if there exists a finite-dimensional subspace H of $L_2(\nu)$ such that $\Pi_H\Theta_1 \cap \Pi_H\Theta_2 = \emptyset$.*

Proof of Lemma 2.2. Let us fix a point $\tau \in \Theta_1$ and consider the problem of testing the hypothesis $S = \tau$ versus $S \in \Theta_2$. By Lemma 2.1, there exists a finite dimensional subspace H_τ such that $\tau \in H_\tau$ and $\tau \notin \Pi_{H_\tau}\Theta_2$. Denote $r_\tau = \rho(\Pi_{H_\tau}\tau, \Pi_{H_\tau}\Theta_2)$ and define the set $U_\tau = \{S : |S - \tau| \leq r_\tau/2,\ \ S \in \Theta_1\}$. It is clear that the sets U_τ and Θ_2 are distinguishable. Since Θ_1 is a compact set, there exists a finite covering of Θ_1 by some sets $U_{\tau_1}, \dots, U_{\tau_n}$ with $\tau_1, \dots, \tau_n \in \Theta_1$. Define the subspace $H = H_{\tau_1} + \ldots + H_{\tau_n}$. Then $\Pi_H\Theta_1 \cap \Pi_H\Theta_2 = \emptyset$, which completes the proof of Lemma 2.2.

In the proof of Theorem 1.1 in the general setting we shall use the following version of Lemma 2.1.

Lemma 2.3. *Let the sets Θ_1 and Θ_2 be closed and bounded in $L_2(\nu)$. Let the hypotheses $\theta \in \Theta_1$ and $\theta \in \Theta_2$ be distinguishable. Then there exists m such that for any $\tau \in \Theta_1$ there exists a finite-dimensional subspace H_τ such that* $\dim(H_\tau) \leq m$ *and $\Pi_{H_\tau}\tau \notin \Pi_{H_\tau}\Theta_2$.*

Proof of Lemma 2.3. Suppose the opposite. Then for any sequence $\rho_n > 0$, $\rho_n \to 0$ as $n \to \infty$ there exists a sequence $S_n \in \Theta_1$ satisfying the following. For every n there exist signals $S_{n1}, \dots, S_{nn} \in \Theta_2$ such that $\sup\{|(S_{nj} - S_n, S_{ni} - S_n)| : 1 \leq i < j \leq n\} < \rho_n$. Arguing similarly to the proof of Lemma 2.1, we come to the contradiction.

Proof of Theorem 1.1. Although the arguments have a rather complicated character, the proof is based basically on the following two facts: Lemma 2.3 and indistinguishability of sets Θ_1 and Θ_2 of the type $\Theta_1 = \bigcup_{i=1}^{\infty}\{\theta_{1i}\}$, $\Theta_2 = \bigcup_{i=1}^{\infty}\{\theta_{2i}\}$, where $\theta_{ti} = (a_{ti1}, a_{ti2}, \dots)$, $a_{tij} = 0$ if $i \neq j$, $t = 1,2$ and $a_{tij} = (-1)^t$ if $i = j$, $t = 1,2$.

Suppose the opposite. Let $\tau_1 \in \Theta_1$ and $\eta_1 \in \Theta_2$. Then by Lemma 2.3 there exist finite dimensional subspaces H_{τ_1}, H_{η_1} such that $\Pi_{H_{\tau_1}}\tau_1 \notin \Pi_{H_{\tau_1}}\Theta_2$ and $\Pi_{H_{\eta_1}}\eta_1 \notin$

$\Pi_{\eta_1}\Theta_1$. We evidently can suppose that $\tau_1 \in H_{\tau_1}$ and $\eta_1 \in H_{\eta_1}$. Denote $H_1 = H_{\tau_1} + H_{\eta_1}$, $\Pi_1 = \Pi_{H_1}$, and let $\Lambda_1 = \Pi_1\Theta_1 \cap \Pi_1\Theta_2 \neq \emptyset$.

Let $\lambda_0 \in \Lambda_1$. There exist the following three possibilities.

i. There exist a vicinity of U_1 of λ_0 and a finite-dimensional subspace H such that

$$\Pi_H(\Pi_1^{-1}U_1 \cap \Theta_1) \cap \Pi_H(\Pi_1^{-1}U_1 \cap \Theta_2) = \emptyset. \tag{2.4}$$

ii. There does not exist finite-dimensional subspace $H \in L_2(\nu)$ such that $\Pi_H(\Pi_1^{-1}\lambda_0 \cap \Theta_1) \cap \Pi_H(\Pi_1^{-1}\lambda_0 \cap \Theta_2) = \emptyset$.

iii. There exist sequences of points $\lambda_i \in H_1$ converging to λ_0 as $i \to \infty$, the vicinities U_i of λ_i and finite-dimensional subspaces H_i, $H_{i-1} \subset H_i$, $\dim(H_i) \to \infty$ as $i \to \infty$, such that

$$\Pi_{H_i}(\Pi_1^{-1}U_j \cap \Theta_1) \cap \Pi_{H_i}(\Pi_1^{-1}U_j \cap \Theta_2) = \emptyset. \quad \textit{if} \quad j \le i \tag{2.5}$$

and

$$\Pi_{H_i}(\Pi_1^{-1}U_j \cap \Theta_1) \cap \Pi_{H_i}(\Pi_1^{-1}U_j \cap \Theta_2) \neq \emptyset \quad \textit{if} \quad j > i. \tag{2.6}$$

It is clear that if for all points $\lambda_0 \in \Lambda_1$ the case i takes place, then the conclusion of Theorem 1.1 is fulfilled.

Consider the case *iii.* By Lemma 2.3, we can define subspaces H_i using the following inductive arguments. Let $z_i \in \Pi_{H_i}(\Pi_1^{-1}U_{i+1} \cap \Theta_1) \cap \Pi_{H_i}(\Pi_1^{-1}U_{i+1} \cap \Theta_2)$ and $\Pi_{H_1}z_i = \lambda_i$. Then there exist $\tau_i \in \Theta_1$ and $\eta_i \in \Theta_2$ such that $\Pi_{H_i}\tau_i = \Pi_{H_i}\eta_i = z_i$. Consider the problems of testing the hypotheses $S = \tau_i$ versus $S \in \Theta_2$ and $S \in \Theta_1$ versus $S = \eta_1$. Applying to these two problems Lemma 2.3, we can find a finite-dimensional subspace $H_{0,i+1}$ such that $\tau_i \in H_{0,i+1}$, $\eta_i \in H_{0,i+1}$ and $\tau_i \notin \Pi_{H_{0,i+1}}\Theta_2$, $\eta_i \notin \Pi_{H_{0,i+1}}\Theta_1$. It is clear that $\Lambda_i = \Pi_{H_i}\Theta_1 \cap \Pi_{H_i}\Theta_2 \neq \emptyset$. We can suppose that the points τ_j, η_j, $j > i$ satisfy the relation $\Pi_{H_i}\tau_j = \Pi_{H_i}\eta_j \in \Lambda_i$ and that there exists a sequence $z_{0i} \in H_i$ such that $\Pi_1 z_{0i} = \lambda_0$ and $\Pi_{H_i}\tau_j \to z_{0i}$ as $j \to \infty$. If the sequences τ_j, η_j do not satisfy this condition we can always pass to subsequences τ_j, η_j possessing this property, using the following diagonal process. Since Λ_2 is a compact we can choose a subsequence $z_{2j} = \Pi_{H_2}\tau_{i_j}$ of the sequence $\Pi_{H_2}\tau_i$ converging to some $z_{02} \in \Lambda_2$. Denoting such a subsequence τ_{i_j} as $\tau_k^{(2)}$, we choose a subsequence $\tau_k^{(3)}$ of $\tau_k^{(2)}$ such that $\Pi_{H_3}\tau_k^{(3)}$ converges to some $z_{03} \in \Lambda_{03}$, and so on. It is clear that the result $\Pi_{H_i}\tau_k^{(k)}$ of the indicated procedure and $\Pi_{H_i}\eta_k^{(k)}$ defined similarly starting with η_i converge to z_{0i} for all i and that $\Pi_{H_1}z_{0i} = \lambda_0$, $\Pi_{H_i}z_{0j} = z_{0i}$ for all $j > i$.

Let us define the sequences of Bayes a priori measures μ_1 and μ_2 on the sets Θ_1 and Θ_2, respectively, with the atoms $\mu_{1m}(z_{1i}) = \frac{1}{m}$, $\mu_{2m}(z_{2i}) = \frac{1}{m}$, $1 \le i \le m$. We denote by π_{tm}, $t = 1, 2$, the corresponding Bayes a posteriori likelihood ratios with respect to the measure of the Gaussian white noise.

We can represent H_i as $H_i = H_{i-1} \oplus H_{ci}$ where $\oplus$ means direct sum. In further arguments we can suppose that $\Pi_{H_i}z_{0j} = \Pi_{H_i}\eta_j = \Pi_{H_i}\tau_j = z_{0i}$ for all $j > i$, since we always can choose a subsequence λ_j such that, for sufficiently large i, the differences $\Pi_{H_i}\tau_j - z_{0i}$, $\Pi_{H_i}\eta_j - z_{0i}$ with $j > i$ are as small as desired. Taking into account this assumption, we choose the system of coordinates in H_i such that the coordinates of z_{0i}, τ_i, η_i are $\Pi_{H_{ci}}z_{0i} = (a_i, 0, \dots, 0)$, $\Pi_{H_{ci}}\tau_i = (a_{1i}, b_{1i}, 0, \dots, 0)$ and $\Pi_{H_{ci}}\eta_i = (a_{1i}, b_{2i}, c_{2i}, 0, \dots, 0)$.

Thus, we can write

$$\pi_{1m} - \pi_{2m} = \frac{1}{m}\sum_{k=1}^{m} J_k \tag{2.7}$$

where

$$J_k = \exp\{\sum_{i=1}^{k-1}\epsilon^{-2}(a_i\psi_i - \frac{1}{2}a_i^2)\}(\exp\{\epsilon^{-2}(a_{1k}\psi_k + b_{1k}\zeta_{1k} - \frac{1}{2}(a_{1k}^2 + b_{1k}^2))\} -$$

$$\exp\{\epsilon^{-2}(a_{2k}\psi_k + b_{2k}\zeta_{1k} + c_{2k}\zeta_{2k} - \frac{1}{2}(a_{2k}^2 + b_{2k}^2 + c_{2k}^2))\}. \tag{2.8}$$

Here ψ_k, ζ_{1k}, ζ_{2k} are independent Gaussian random variables, $E\psi_k = E\zeta_{1k} = E\zeta_{2k} = 0$, $E\psi_k^2 = E\zeta_{1k}^2 = E\zeta_{2k}^2 = \epsilon^2$.

Since $EJ_{k_1}J_{k_2} = 0$ if $k_1 \neq k_2$ then

$$\lim_{m\to\infty} E^2|\pi_{1m} - \pi_{2m}| = \lim_{m\to\infty} E[\pi_{1m} - \pi_{2m}]^2 = 0 \tag{2.9}$$

Since m does not depend on ϵ, we come to the contradiction.

In the case *ii* the arguments are the same as in *iii*. The only difference is that now we set $\lambda_i = \lambda_0$. The proof of Theorem 1.1 is completed.

3. Proof of Theorem 1.1 for the case of testing hypotheses about density. The arguments are basically the same as in the case of signal detection, up to evident modification of terminology and some changes in analytical relations.

Let us first outline the modifications of the corresponding version of Lemma 2.1. As a consequence, we obtain also the results of Lemmas 2.2 and 2.3.

Proof of Lemma 2.1 for the case of testing hypotheses about density. We can preserve the same arguments as in the case of signal detection; all we need in addition is to suppose that $|(S_0, S_k - S_0)| < \rho_k$. For all j denote $\phi_j = S_j - S_0$. For all j_1, j_2 we set $r_{j_1 j_2} = (\phi_{j_1}, \phi_{j_2})$, $r_{0j_1} = (S_0, \phi_{j_1})$ and $r_{00} = (S_0, S_0)$.

The corresponding likelihood ratio (compare with (2.1)) w.r.t. measure ν equals

$$I_0 = \prod_{s=1}^{m} S_0(X_s), \qquad I_m = \frac{1}{m}\sum_{j=1}^{m}\prod_{s=1}^{n} S_j(X_s) = \frac{1}{m}\sum_{j=1}^{m}\prod_{s=1}^{n}(S_0(X_s) + \phi_j(X_s)) \tag{3.1}$$

in the case of the hypothesis and its Bayes alternative, respectively.

By direct calculation we get

$$E_\nu[I_m - I_0]^2 = D_{m1} - D_{m2}, \quad D_{m1} = D_{m11} + D_{m12} \tag{3.2}$$

where

$$D_{m11} = 2m^{-2}\sum_{1\le j_1<j_2\le m}(r_{00} + r_{0j_1} + r_{0j_2} + r_{j_1j_2})^n - \frac{m-1}{m}r_{00}^n, \tag{3.3}$$

$$D_{m12} = m^{-2}\sum_{j=1}^{k}(r_{00} + 2r_{0j} + r_{jj})^n - m^{-1}r_{00}^n =$$

$$m^{-2}\sum_{j=1}^{k}||S_j||^{2n} - m^{-1}||S_0||^{2n}, \tag{3.4}$$

$$D_{m2} = 2m^{-1}\sum_{j=1}^{k}(r_{00} + r_{0j})^n - 2r_{00}^n \tag{3.5}$$

Since the choice of m and of the sequence ρ_k does not depend on n, (3.2)-(3.5) together imply (2.2). This completes the proof of Lemma 2.1.

Proof of Theorem 1.1 for the case of testing hypotheses about density. We can use the same arguments as for the problem of signal detection, till the definition of the likelihood ratios π_{1m} and π_{2m}. The likelihood ratios π_{1m} and π_{2m} now are taken w.r.t. measure ν and are

$$\pi_{1m} = \frac{1}{m}\sum_{k=1}^{m} J_{1k}, \qquad \pi_{2m} = \frac{1}{m}\sum_{k=1}^{m} J_{2k} \tag{3.6}$$

with

$$J_{1k} = \prod_{s=1}^{n}(\lambda_0(X_s) + \sum_{i=1}^{k-1} a_i\psi_i(X_s) + a_{1k}\psi_k(X_s) + b_{1k}\zeta_{1k}(X_s)),$$

$$J_{2k} = \prod_{s=1}^{n}(\lambda_0(X_s) + \sum_{i=1}^{k-1} a_i\psi_i(X_s) + a_{2k}\psi_k(X_s) + b_{2k}\zeta_{1k}(X_s) + c_{2k}\zeta_{2k}(X_s)).$$

Here $\psi_1, \dots, \psi_k$, ζ_{1k}, ζ_{2k} are the orthonormal basis orths of properly chosen coordinates (compare with (2.8)).

Denote

$$g_k(x) = \lambda_0(x) + \sum_{i=1}^{k-1} a_i\psi_i(x),$$

$$\phi_{1k}(x) = a_{1k}\psi_k(x) + b_{1k}\zeta_{1k}(x),$$

$$\phi_{2k}(x) = a_{2k}\psi_k(x) + b_{2k}\zeta_{1k}(x) + c_{2k}\zeta_{2k}(x).$$

Then, for t=1,2,

$$J_{tk} = \prod_{s=1}^{n}(g_k(X_s) + \phi_{tk}(X_s)).$$

For any $k_1 < k_2$ and $i,j = 1,2$ we have

$$(g_{k_1} + \phi_{ik_1}, g_{k_2} + \phi_{jk_2}) = ||g_{k_1}||^2 + (\phi_{ik_1}, g_{k_2}). \tag{3.7}$$

This implies $E_\nu[J_{ik_1}(J_{1k_2} - J_{2k_2})] = 0$. Hence $E_\nu[(J_{1k_1} - J_{2k_1})(J_{1k_2} - J_{2k_2})] = 0$.

Therefore

$$E_\nu[\pi_{1m} - \pi_{2m}]^2 = m^{-2}\sum_{k=1}^{m}(||g_k + \phi_{1k}||^{2n} +$$

$$||g_k + \phi_{2k}||^{2n} - 2(g_k + \phi_{1k}, g_k + \phi_{2k})^{2n}). \tag{3.8}$$

Hence

$$\lim_{m\to\infty} E_\nu[\pi_{1m} - \pi_{2m}]^2 = 0. \tag{3.9}$$

This completes the proof of Theorem 1.1.

References

1. Brown, L.D. and Low, M.G. (1992) Asymptotic equivalence of nonparametric regression and white noise. To appear Ann. Statist.
2. Burnashev, M.V. (1979) On the minimax solution of inaccurately known signal in a white Gaussian noise. Background. Theory Probab. Appl., 24 107-119
3. Donoho, D.L. and Liu, R.C. (1987) Geometrizing rates of convergence, I. Technical Report 137a, Department of Statistics, University of California, Berkeley.
4. Ermakov, M.S. (1990) Minimax detection of signal in a Gaussian white noise. Theory Probab. Appl., 35 704-715.
5. Ermakov, M.S. (1995) On asymptotic minimaxity of tests of Kolmogorov and omega square type. Theory Probab. Appl. 40 54-67.
6. Ibragimov, I.A. and Khasminskii, R.Z. (1977) On estimation of infinitely dimensional parameter in Gaussian white noise. Doklady AN USSR 236 1053-1055.
7. Ingster, Yu.I. (1988) Minimax testing of simple hypothesis versus weakly nonparametric sets of alternatives. Theory Probab. Appl. 33 592-596.
8. Ingster, Yu.I. (1993) Asymptotically minimax hypothesis testing for nonparametric alternatives. Mathematical Methods of Statistics. 2 86-113, 171-189, 249-268.
9. Mann, H.B. and Wald, A. (1942) On the choice of the number of intervals in the application of the chi-square tests. Ann. Math. Statist. 13 306-317.
10. Nussbaum, M. (1995) Asymptotic equivalence of density estimation and white noise. Preprint No 174 Weierstrass-Institute fur Angewandte Analysis and Stochastik. Berlin.
11. Stein, C. (1956) Efficient nonparametric testing and estimation. Third Berkeley Symp. Math. Statist. and Probab. 1 Univ. California Press Berkeley 187-195.

Recent publications of the
Weierstraß–Institut für Angewandte Analysis und Stochastik

Preprints 1996

244. Grigori N. Milstein: Stability index for invariant manifolds of stochastic systems.

245. Luis Barreira, Yakov Pesin, Jörg Schmeling: Dimension of hyperbolic measures – A proof of the Eckmann–Ruelle conjecture.

246. Leonid M. Fridman, Rainer J. Rumpel: On the asymptotic analysis of singularly perturbed systems with sliding mode.

247. Björn Sandstede: Instability of localised buckling modes in a one–dimensional strut model.

248. Björn Sandstede, Christopher K.R.T. Jones, James C. Alexander: Existence and stability of N–pulses on optical fibers with phase–sensitive amplifiers.

249. Vladimir Maz'ya, Gunther Schmidt: Approximate wavelets and the approximation of pseudodifferential operators.

250. Gottfried Bruckner, Sybille Handrock–Meyer, Hartmut Langmach: On the identification of soil transmissivity from measurements of the groundwater level.

251. Michael Schwarz: Phase transitions of shape memory alloys in soft and hard loading devices.

252. Gottfried Bruckner, Masahiro Yamamoto: On the determination of point sources by boundary observations: uniqueness, stability and reconstruction.

253. Anton Bovier, Véronique Gayrard: Hopfield models as generalized random mean field models.

254. Matthias Löwe: On the storage capacity of the Hopfield model.

255. Grigori N. Milstein: Random walk for elliptic equations and boundary layer.

256. Lutz Recke, Daniela Peterhof: Abstract forced symmetry breaking.

257. Lutz Recke, Daniela Peterhof: Forced frequency locking in $\mathbf{S}^1$–equivariant differential equations.

258. Udo Krause: Idealkristalle als Abelsche Varietäten.

259. Nikolaus Bubner, Jürgen Sprekels: Optimal control of martensitic phase transitions in a deformation–driven experiment on shape memory alloys.

260. Christof Külske: Metastates in disordered mean field models: random field and Hopfield models.

261. Donald A. Dawson, Klaus Fleischmann: Longtime behavior of a branching process controlled by branching catalysts.

262. Tino Michael, Jürgen Borchardt: Convergence criteria for waveform iteration methods applied to partitioned DAE systems in chemical process simulation.

263. Michael H. Neumann, Jens–Peter Kreiss: Bootstrap confidence bands for the autoregression function.

264. Silvia Caprino, Mario Pulvirenti, Wolfgang Wagner: Stationary particle systems approximating stationary solutions to the Boltzmann equation.

265. Wolfgang Dahmen, Angela Kunoth, Karsten Urban: Biorthogonal spline–wavelets on the interval – Stability and moment conditions.

266. Daniela Peterhof, Björn Sandstede, Arnd Scheel: Exponential dichotomies for solitary–wave solutions of semilinear elliptic equations on infinite cylinders.

267. Andreas Rathsfeld: A wavelet algorithm for the solution of a singular integral equation over a smooth two–dimensional manifold.

268. Jörg Schmeling, Serge E. Troubetzkoy: Dimension and invertibility of hyperbolic endomorphisms with singularities.

269. Erwin Bolthausen, Dmitry Ioffe: Harmonic crystal on the wall: a microscopic approach.

270. Nikolai N. Nefedov, Klaus R. Schneider: Delayed exchange of stabilities in singularly perturbed systems.

271. Michael S. Ermakov: On large and moderate large deviations of empirical bootstrap measure.

272. Priscilla E. Greenwood, Jiaming Sun: On criticality for competing influences of boundary and external field in the Ising model.