The Annals of Statistics
1998, Vol. 26, No. 1, 279–287

ASYMPTOTIC NONEQUIVALENCE OF NONPARAMETRIC EXPERIMENTS WHEN THE SMOOTHNESS INDEX IS 1/2

BY LAWRENCE D. BROWN[1] AND CUN-HUI ZHANG[2]

University of Pennsylvania and Rutgers University

An example is provided to show that the natural asymptotic equivalence does not hold between any pairs of three nonparametric experiments: density problem, white noise with drift and nonparametric regression, when the smoothness index of the unknown nonparametric function class is 1/2.

1. Introduction. There have recently been several papers demonstrating the global asymptotic equivalence of certain nonparametric problems. See especially Brown and Low (1996), who established global asymptotic equivalence of the usual white-noise-with-drift problem to the nonparametric regression problem, and Nussbaum (1996), who established global asymptotic equivalence to the nonparametric density problem. In both these instances the results were established under a smoothness assumption on the unknown nonparametric drift, regression or density function. In both cases such functions were assumed to have smoothness coefficient $\alpha > 1/2$, for example, to satisfy

$$|f(x) - f(y)| \le M|x - y|^{\alpha} \tag{1.1}$$

for all (x, y) in their domain of definition.

This note contains an example which shows that such a condition is necessary, in the sense that global asymptotic equivalence may fail between any pairs of the above three nonparametric experiments when (1.1) fails in a manner that the nonparametric family of unknown functions contains functions satisfying (1.1) with $\alpha = 1/2$ but not with any $\alpha > 1/2$.

Efromovich and Samarov (1996) have already shown that asymptotic equivalence of nonparametric regression and white noise may fail when $\alpha < 1/4$ in (1.1). The present counterexample to equivalence is somewhat different from theirs and carries the boundary value $\alpha = 1/2$.

Section 2 contains a brief formal description of the nonparametric problems and of global asymptotic equivalence. Section 3 describes the example.

Received July 1996; revised May 1997.
[1]Supported in part by NSF Grant DMS-91-07842.
[2]Supported in part by the National Security Agency.
AMS 1991 *subject classifications*. Primary 62G07; secondary 62G20, 62M05.
Key words and phrases. Risk equivalence, nonparametric regression, density estimation, white noise.

2. The problem.

Density problem $\xi_{0,n}$. In the nonparametric density problem one observes i.i.d. variables $X_1, \dots, X_n$ having some density $g = f^2/4$. Assume the support of f is $[0, 1]$. About f it is assumed only that $f \in \mathscr{F} \cap \{f: f \geq 0, \int_0^1 f^2(x)\,dx = 4\}$, where $\mathscr{F}$ is some (large) class of functions in $L_2[0, 1]$. In general, the goal is to use the observations to make some sort of inference about f.

White noise $\xi_{1,n}$. In the white-noise problem one observes a Gaussian process $\{Z_n(t),\ 0 \leq t \leq 1\}$ which can be symbolically written as

$$dZ_n(t) = f(t)\,dt + \frac{1}{\sqrt{n}}\,dB(t),$$

where $B(t)$ denotes the standard Brownian motion on $[0, 1]$. Again $f \in \mathscr{F}$.

Nonparametric regression $\xi_{2,n}$. Here one observes (Y_i, X_i), $1 \leq i \leq n$. Given $\{X_i: 1 \leq i \leq n\}$, the Y_i are independent normal variables with mean $f(X_i)$ and unit variance, $f \in \mathscr{F}$. For deterministic X_i [e.g., $X_i = i/(n+1)$] and $\alpha = 1/2$, the asymptotic nonequivalence of nonparametric regression and white noise has already been established in Brown and Low [(1996), Remark 4.6]. Hence, in the case of current interest the X_i are i.i.d. uniform random variables on $[0, 1]$.

Asymptotic equivalence. The assertion that two of these formulations—say, nonparametric regression and white noise—are globally asymptotically equivalent is equivalent to the following assertion: for each n, let A_n be an action space and L_n be a loss function with $\|L_n\|_\infty \leq 1$, and let δ_n be a procedure in one of the problems. Then there exists a corresponding procedure δ_n' in the other problem such that the sequences δ_n and δ_n' are asymptotically equivalent, which means

$$\lim_{n\to\infty} \sup_{f\in\mathscr{F}} \left| E_f^{(n)}\big(L_n(f, \delta_n)\big) - E_f^{(n)'}\big(L_n(f, \delta_n')\big) \right| = 0. \tag{2.1}$$

[The expectations in (2.1) are computed under f for the nth form of the first and second problems, respectively. The notation should be interpreted to allow for randomized decision rules. The convergence in (2.1) is uniform in the loss functions and decision rules as they are allowed to change with n.] The asymptotic equivalence established in Brown and Low (1996) is a little different from the above statement. The equivalence expressed above is established in Brown and Zhang (1996).

The equivalence assertion involving the density problem was established by Nussbaum (1996) under (1.1) and the additional assumption that $\min_{0\leq x\leq 1} f(x)$ is uniformly bounded away from 0 over $\mathscr{F}$: it is that the density problem with unknown density $g = f^2/4$ is asymptotically equivalent to the white noise with drift f. The statement of (2.1) thus becomes

$$\lim_{n\to\infty} \sup_{f\in\mathscr{F}} \left| E_{f^2/4}^{(n)}\big(L_n(f, \delta_n)\big) - E_f^{(n)'}\big(L_n(f, \delta_n')\big)\right| = 0,$$

where the first expectation refers to the density problem and the second to either the white-noise or the nonparametric regression problem.

Hence, in order to prove two formulations nonequivalent in this sense, it suffices to find some action spaces A_n, uniformly bounded loss functions L_n and priors on $\mathscr{F}$ for which the Bayes risks converge to different values in different problems.

3. The example. Let $\mathscr{F}_{\alpha, M}$ be the class of all functions f with support $[0, 1]$ such that (1.1) holds for all $0 \le x < y \le 1$. Here we shall give an example to show that the white-noise and nonparametric regression experiments are not asymptotically equivalent for $\mathscr{F} = \mathscr{F}_{1/2, M}$, and also that under the stronger restriction $\mathscr{F} = \mathscr{F}_{1/2, M} \cap \{f\colon f > \varepsilon_0, \|f\|_2^2 = 4\}$, the density problem is not asymptotically equivalent to either the white noise or the nonparametric regression for every $0 \le \epsilon_0 < 2$.

Let $\psi(x)$ be a function in $\mathscr{F}_{1, M}$ such that $\int_0^1 \psi(x)\,dx = \int_0^1 \psi^3(x)\,dx = 0$ and $\psi(x) = 0$ for $x \notin (0, 1)$. For $m = m_n$, define

$$\phi_j(x) = m^{-1/2}\psi(m\{x - (j-1)/m\}), \qquad 1 \le j \le m. \tag{3.1}$$

For $\theta = (\theta_1, \dots, \theta_m)$ with $|\theta_j| \le 1$, define

$$\phi_\theta(x) = \sum_{j=1}^m \theta_j \phi_j(x), \qquad f_\theta(x) = 2 + \phi_\theta(x) - c_m, \tag{3.2}$$

where $c_m = c_m(\theta)$ is chosen so that $\int (f_\theta/2)^2 = 1$.

MOTIVATION. Examples of asymptotic nonequivalence can be constructed by finding Bayes problems for which the Bayes risks converge to different limits for different sequences of experiments. Consider prior distributions on the subspaces $\{m^{1/2-\alpha}\phi_\theta\}$ of $\mathscr{F}_{\alpha, M}$ with ϕ_θ in (3.2) and $m \sim n$. Due to the normality of the errors, the nonparametric regression $\xi_{2, n}$ is characterized by the Fisher information $\sigma_{\alpha, j}^2 = m^{1-2\alpha}\sum_i \phi_j^2(X_i)$ given $\{X_i\}$, $1 \le j \le m$, so that its Bayes risk is of the form $Er_n(\sigma_{\alpha, 1}, \dots, \sigma_{\alpha, n})$, with r_n being the conditional Bayes risk. It is easily seen that the corresponding Bayes risk for the white noise $\xi_{1, n}$ is $r_n(\sqrt{E\sigma_{\alpha, 1}^2}, \dots, \sqrt{E\sigma_{\alpha, n}^2})$. For tractable Bayes problems, the limit Bayes risk is often found via Taylor expansions of certain components of r_n and applications of limiting theorems to $\sum_j \sigma_{\alpha, j}^k$ for $\xi_{2, n}$. Since $\sigma_{\alpha, j}$ is proportional to $m^{-\alpha}$, higher-order terms are needed in Taylor expansions only for small α, whereas lower-order terms are more likely to be crucial for large α. Since we are interested in the largest possible $\alpha = 1/2$ for nonequivalence examples, we shall look for Bayes problems in which $\sum_j \sigma_{\alpha, j}$ plays an important role. This leads to the following variation of the compound hypothesis testing problem of Robbins (1951).

Let G be the prior such that θ_j, $1 \le j \le m$, are i.i.d. Bernoulli variables with

$$P_G(\theta_j = 1) = P_G(\theta_j = -1) = 1/2. \tag{3.3}$$

Consider the estimation of $\theta = (\theta_1, \dots, \theta_m)$ with the 0–1 loss for picking more than half of the θ_j wrong,

$$L(\theta, a) = I\left\{\sum_{j=1}^{m} I\{\theta_j \ne a_j\} > m/2\right\}, \tag{3.4}$$

where $a = (a_1, \dots, a_m)$ is the action. We shall show that the Bayes risks for the three sequences of experiments converge to different limits as $n \to \infty$ and $n/m \to \lambda$,

$$R(G, d^G;\ \xi_{j,n}) \to \Phi(-\tau_j(\lambda, \psi)), \tag{3.5}$$

where $\{\xi_{j,n}\}$, $j = 0, 1, 2$, are, respectively, the density problem, white noise and nonparametric regression, $d^G = d^G(\xi_{j,n})$ is the Bayes rule with experiment $\xi_{j,n}$, $\Phi(\cdot)$ is the standard normal distribution function and τ_j is defined below. Let N be a Poisson variable with $EN = \lambda$ and let U_i, $i \ge 1$, be i.i.d. uniform $(0, 1)$ variables independent of N. The functions $\tau_j(\lambda, \psi)$ in (3.5) are analytically different and are given by

$$\begin{aligned} \tau_0(\lambda, \psi) &= E\left|\sum_{i=1}^{N} \psi(U_i)\right|, \\ \tau_1(\lambda, \psi) &= \sqrt{\frac{2\lambda}{\pi}}\|\psi\|_2, \\ \tau_2(\lambda, \psi) &= E\sqrt{\frac{2}{\pi}\sum_{i=1}^{N} \psi^2(U_i)}. \end{aligned} \tag{3.6}$$

By the Schwarz inequality, $\tau_2(\lambda, \psi) < \tau_1(\lambda, \psi)$. For small λ,

$$\tau_1(\lambda, \psi) > \tau_0(\lambda, \psi) = (\|\psi\|_1 + o(1))\lambda > \tau_2(\lambda, \psi) = \sqrt{2/\pi}\,(\|\psi\|_1 + o(1))\lambda.$$

[By the moment convergence in the strong law of large numbers and the central limit theorem, $\tau_i(\lambda, \psi)/\tau_j(\lambda, \psi) \to 1$ as $\lambda \to \infty$ for all $0 \le i < j \le 2$.]

For $\mathscr{F} = \mathscr{F}_{1/2, M}$, the asymptotic nonequivalence between the white noise and the nonparametric regression follows from (3.5) and (3.6) as $\psi \in \mathscr{F}_{1, M}$ implies $f_\theta \in \mathscr{F}_{1/2, M}$ in view of (3.1) and (3.2). Since $\phi_j(x)\phi_k(x) = 0$ for $j \ne k$,

$$\|\phi_\theta\|_2^2 = m\|\phi_j\|_2^2 = m^{-1}\|\psi\|_2^2, \qquad \|\phi_\theta\|_\infty = \|\psi\|_\infty/\sqrt{m}. \tag{3.7}$$

Since $g_\theta = (f_\theta/2)^2$ is a density, by (3.7) c_m do not depend on θ under (3.3) and

$$c_m = 2\left\{1 - \sqrt{1 - \|\phi_\theta\|_2^2/4}\right\} = m^{-1}\|\psi\|_2^2/4 + O(m^{-2}). \tag{3.8}$$

The asymptotic nonequivalence between the density problem and either the white noise or the nonparametric regression also follows from (3.5) and (3.6) as the further restrictions $\|f_\theta\|_2^2 = 4$ and $f_\theta > \epsilon_0$ hold for large m and fixed $0 \le \epsilon_0 < 2$ by (3.2), (3.7), (3.8) and the fact that $\|\psi\|_\infty \le M$.

MOTIVATION (continued). The decision problem (3.4) is closely related to compound hypothesis testing and estimation with L_1 loss. In the compound hypothesis testing problem, the loss function is $m^{-1}\sum_{j=1}^m I\{\theta_j \ne a_j\}$, the Bayes rules for the prior (3.3) are the same as with the loss (3.4), and the Bayes risks are $1/2-\{\lambda/(4n)\}^{1/2}\tau_j(\lambda, \psi)+o(n^{-1/2})$ for $\xi_{j,n}$. In the estimation problem with the L_1 loss $\int_0^1 |a(t) - f(t)|\,dt$, the Bayes rules are $f^G(\cdot) = \sum_{j=1}^m d_j^G \phi_j(\cdot)$ with ϕ_j in (3.1) and $d^G = (d_1^G, \dots, d_m^G)$ in (3.5), and the Bayes risks are

$$\sqrt{\lambda/n}\|\psi\|_1\big[1/2 - \{\lambda/(4n)\}^{1/2}\tau_j(\lambda, \psi)\big] + o(n^{-1}).$$

Thus, the difference among the nonparametric experiments is recovered in the second-order asymptotics in the above two decision problems. The proofs of the above statements are omitted as they are similar to and simpler than the calculation of the Bayes risks (3.5) and (3.6).

REMARK 1. In nonparametric regression with deterministic $X_i = i/(n+1)$, Y_i contain no information about f in the subspace (3.2) with $f = f_\theta$ and $m = n+1$ as in Brown and Low (1996), so that the Bayes risk for the decision problem (3.4) is 1/2 under the prior (3.3). Since $\tau_j(1, \psi) > 0$ for $\|\psi\|_2 > 0$ in (3.5) for each $j = 0, 1, 2$, the nonparametric regression with deterministic $\{X_i\}$ is asymptotically nonequivalent to $\{\xi_{j,n}\}$ for $\alpha = 1/2$.

REMARK 2. Many applications of the nonparametric experiments discussed here involve their d-dimensional versions with f being an unknown function of d real variables (e.g., $E[Y_i|X_i] = f(X_i)$ with X_i being uniform $[0, 1]^d$ in the case of nonparametric regression). The above example can be easily modified to show the asymptotic nonequivalence of these nonparametric experiments when the smoothness index is $d/2$.

4. Calculation of Bayes risks. In this section we calculate the limit of the Bayes risks given in (3.5) and (3.6).

Nonparametric regression $\xi_{2,n}$. Let P_n^* be the conditional probability given $X_1, \dots, X_n$ and under P_G. Set

$$S_j = \sum_{i=1}^n \phi_j(X_i)(Y_i - 2 + c_m), \qquad \sigma_j^2 = \sum_{i=1}^n \phi_j^2(X_i).$$

Since $\phi_j(\cdot)$ have disjoint support sets, by (3.2) S_j are sufficient for θ_j. In addition, (S_j, θ_j), $1 \le j \le m$, are independent random vectors under P_n^* with

$$P_n^*\{S_j \le t|\theta_j\} = \Phi\big((t - \theta_j\sigma_j^2)/\sigma_j\big), \qquad P_n^*\{\theta_j = \pm 1\} = 1/2.$$

Since the loss function in (3.4) is increasing in $I\{\theta_j \neq a_j\}$, the Bayes rule $d^G = (d_1^G, \dots, d_m^G)$ is given by

$$d_j^G = I\{S_j > 0\} - I\{S_j \leq 0\},$$

and $I\{\theta_j \neq d_j^G\}$ are independent Bernoulli random variables under P_n^* with

$$p_j = P_n^*\{\theta_j \neq d_j^G\} = P_n^*\{\theta_j S_j < 0 | \theta_j\} = \Phi(-\sigma_j). \tag{4.1}$$

Since $\sigma_j^2 \leq n\|\phi_j\|_\infty^2 = (n/m)\|\psi\|_\infty^2 = O(1)$, $p_j(1-p_j)$ are uniformly bounded away from zero, so that

$$\left(\sum_{j=1}^m p_j(1-p_j)\right)^{-1/2}\left\{\sum_{j=1}^m I\{\theta_j \neq d_j^G\} - \sum_{j=1}^m p_j\right\} \to_{\mathscr{D}} N(0,1) \tag{4.2}$$

uniformly under P_n^*. Let $N_j = \#\{i\colon (j-1)/m < X_i \leq j/m\}$. Since X_i are i.i.d. uniform, $(N_1, \dots, N_m)$ is a multinomial vector with $EN_j = n/m \to \lambda$. It follows that

$$\sqrt{m}E\sigma_1 = E\left\{\sum_{i=1}^{N_1}\psi^2(U_i)\right\}^{1/2} \to \tau_2(\lambda, \psi)\sqrt{\frac{\pi}{2}},$$

$$mE\sigma_1^2 = mn\|\phi_1\|_2^2 = \frac{n}{m}\|\psi\|_2^2$$

and

$$mE\sigma_1\sigma_2 = E\left\{\sum_{i=1}^{N_1}\psi^2(U_i)\sum_{i=N_1+1}^{N_2}\psi^2(U_i)\right\}^{1/2} \to \tau_2^2(\lambda, \psi)\frac{\pi}{2}$$

as $n \to \infty$ and $n/m \to \lambda$, where $\{U_i\}$ are i.i.d. uniform $(0,1)$ variables independent of $\{X_i\}$. These lead to $mE\sigma_1^2 = O(1)$ and

$$\operatorname{Var}\left(\sum_{j=1}^m \sigma_j\right) \leq mE\sigma_1^2 + m(m-1)\{E\sigma_1\sigma_2 - (E\sigma_1)^2\} = o(m).$$

By (4.1) and the boundedness and Taylor expansion of $\Phi(\cdot)$, $p_j - 1/2 = \Phi(-\sigma_j) - 1/2 = -\sigma_j/\sqrt{2\pi} + O(\sigma_j^2)$ and $p_j(1-p_j) = 1/4 - (p_j - 1/2)^2 = 1/4 + O(\sigma_j^2)$ with $O(1)$ uniformly bounded by a universal constant, so that by the Chebyshev inequality,

$$\frac{\sum_{j=1}^m (p_j - 1/2)}{\sqrt{m/4}} \to -\tau_2(\lambda, \psi) \quad \text{and} \quad \frac{\sum_{j=1}^m p_j(1-p_j)}{m} \to \frac{1}{4}$$

in probability. This and (4.2) imply (3.5) and (3.6) for $\xi_{2,n}$, as

$$P\left\{\sum_{j=1}^m I\{\theta_j \neq d_j^G\} > m/2\right\} = E\Phi\left(\left(\sum_{j=1}^m p_j(1-p_j)\right)^{-1/2}\sum_{j=1}^m\left(p_j - \tfrac{1}{2}\right)\right) + o(1)$$

$$= \Phi(-\tau_2(\lambda, \psi)) + o(1).$$

White noise $\xi_{1,n}$. The calculation is similar but much simpler compared with nonparametric regression. The sufficient statistics are

$$Z_j = n\int \phi_j(t)\,dZ_n(t) \sim N(\theta_j n\|\phi_j\|_2^2, n\|\phi_j\|_2^2),$$

so that the Bayes rule $d^G = (d_1^G, \dots, d_m^G)$ is given by $d_j^G = I\{Z_j > 0\} - I\{Z_j \le 0\}$. Consequently, $I\{\theta_j \neq d_j^G\}$, $1 \le j \le m$, are i.i.d. Bernoulli variables with $P\{\theta_j \neq d_j^G\} = P\{\theta_j \neq d_j^G | \theta_j\} = \Phi(-\sqrt{n}\|\phi_j\|_2)$. Since $\sqrt{n}\|\phi_j\|_2 = \sqrt{n}\|\psi\|_2/m = (\sqrt{\lambda} + o(1))\|\psi\|_2/\sqrt{m}$ and $\Phi(t) \sim 1/2 + t/\sqrt{2\pi}$ for small t,

$$P\left\{\sum_{j=1}^m I\{\theta_j \neq d_j^G\} > m/2\right\} \to \Phi(-\|\psi\|_2\sqrt{2\lambda/\pi}).$$

Density problem $\xi_{0,n}$. For $1 \le j \le m$, define

$$\Lambda_j(\pm 1) = \exp\left[2\sum_{i=1}^n \log\left(\frac{1 \pm \phi_j(X_i)}{2} - \frac{c_m}{2}\right) I\left\{\frac{j-1}{m} < X_i \le \frac{j}{m}\right\}\right].$$

Since the observations $X_1, \dots, X_n$ are i.i.d. from $g_\theta = (f_\theta/2)^2$, by (3.2) the likelihood is

$$\prod_{i=1}^n g_\theta(X_i) = \prod_{j=1}^m \Lambda_j(\theta_j),$$

so that θ_j, $1 \le j \le m$, are independent given $X_1, \dots, X_n$ and Λ_j (± 1) are sufficient for θ_j. Consequently, the Bayes rule $d^G = (d_1^G, \dots, d_m^G)$ is given by

$$d_j^G = I\{\Lambda_j(1) > \Lambda_j(-1)\} - I\{\Lambda_j(1) \le \Lambda_j(-1)\},$$

and $I\{\theta_j \neq d_j^G\}$, $1 \le j \le m$, are independent variables given $X_1, \dots, X_n$ with

$$p_j = P_n^*\{\theta_j \neq d_j^G\} = P_n^*\{\theta_j \neq d_j^G | \theta_j\} = \frac{\min(\Lambda_j(1), \Lambda_j(-1))}{\Lambda_j(1) + \Lambda_j(-1)}, \tag{4.3}$$

where P_n^* is the conditional probability given $X_1, \dots, X_n$ (the posterior probability measure with respect to θ). Taking Taylor expansions, we find by (3.7) and (3.8),

$$2\log(1 \pm \phi_j(X_i)/2 - c_m/2) = \pm\phi_j(X_i) - c_m - \phi_j^2(X_i)/4 + O(m^{-3/2})$$

with uniform $O(1)$, so that

$$\log\{\Lambda_j(1)/\Lambda_j(-1)\} = 2\tilde{S}_j + O(N_j m^{-3/2}), \tag{4.4}$$

where $\tilde{S}_j = \sum_{i=1}^n \phi_j(X_i)$. Since $\int_0^1 \psi(x)\,dx = \int_0^1 \psi^3(x)\,dx = 0$,

$$\begin{aligned} E_\theta \phi_j(X_i) &= \int \phi_j(x)(1 + \theta_j\phi_j(x)/2 - c_m/2)^2\,dx \\ &= (1 - c_m/2)\theta_j\|\phi_j\|_2^2 \\ &= (1 - c_m/2)\theta_j\|\psi\|_2^2/m^2, \end{aligned}$$

and by (3.7) and (3.8),

$$\begin{aligned} E_\theta \phi_j^2(X_i) &= (1 - c_m/2)^2 \|\phi_j\|_2^2 + \|\phi_j\|_4^4/4 \\ &\le (1 - \|\psi\|_2^2/(4m))\|\psi\|_2^2/m^2 + \|\psi\|_4^4/(4m^3) \\ &\le \|\psi\|_\infty^2/m^2. \end{aligned}$$

Since $\|\phi_j(x)\|_\infty = \|\psi\|_\infty/\sqrt{m}$, by the Bernstein inequality there exists a constant C such that

$$P_\theta\left\{|\tilde{S}_j| > \frac{(C\log m + 1)}{\sqrt{m}}\right\} \le \frac{1}{m^2}$$

uniformly in θ, so that

$$P\left\{\max_{1\le j\le m} |\tilde{S}_j| > \frac{(C\log m + 1)}{\sqrt{m}}\right\} \le \frac{1}{m} \to 0.$$

This and (4.4) allow us to take the Taylor expansion of p_j in (4.3),

$$\begin{aligned} 1 - 2p_j &= \frac{|\Lambda_j(1)/\Lambda_j(-1) - 1|}{\{\Lambda_j(1)/\Lambda_j(-1) + 1\}} \\ &= \frac{|2\tilde{S}_j + 2\tilde{S}_j^2|}{2 + 2\tilde{S}_j} + O(m^{-3/2})\{N_j + (\log m)^3\} \qquad (4.5) \\ &= |\tilde{S}_j| + O(m^{-3/2})\{N_j + (\log m)^3\}. \end{aligned}$$

Since $\sum_{j=1}^m N_j = n = O(m)$, $\max_{1\le j\le m} |\frac{1}{2} - p_j| \to 0$ in probability, so that (4.2) holds under the current P_n^* over some events $C_n \in \sigma(X_1, \dots, X_n)$ with $P\{C_n\} \to 1$. Thus

$$R(G, d^G;\ \xi_{0,n}) = E\Phi\left(\left(\sum_{j=1}^m p_j(1 - p_j)\right)^{-1/2} \sum_{j=1}^m (p_j - 1/2)\right) + o(1). \qquad (4.6)$$

In addition,

$$\sqrt{m}E_\theta|\tilde{S}_1| = E_\theta\left|\sum_{i=1}^{N_1} \psi(\tilde{U}_{i1})\right|,$$

$$mE_\theta|\tilde{S}_1\tilde{S}_2| = E_\theta\left|\sum_{i=1}^{N_1} \psi(\tilde{U}_{i1}) \sum_{i=1}^{N_2} \psi(\tilde{U}_{i2})\right|,$$

where $\tilde{U}_{ij}$, $i \ge 1$, are i.i.d. with density $(1 + \theta_j m^{-1/2}\psi(x)/2 - c_m/2)^2$ under P_θ. Since these density functions converge uniformly in (x, θ_j) to the uniform $(0, 1)$ density and $(N_1, \dots, N_m)$ is a multinomial vector with $EN_j = E_\theta N_j = n/m$,

$$\sqrt{m}E|\tilde{S}_1| \to E\left|\sum_{i=1}^{N} \psi(U_i)\right| = \tau_0(\lambda, \psi), \qquad mE|\tilde{S}_1\tilde{S}_2| \to \tau_0^2(\lambda, \psi),$$

and $E\tilde{S}_1^2 \le nE\phi_j^2(X_1) \le n\|\psi\|_\infty^2/m^2$, so that

$$E\left|\frac{1}{\sqrt{m}}\sum_{j=1}^m |\tilde{S}_j| - \tau_0(\lambda, \psi)\right|^2 \to 0, \qquad E\sum_{j=1}^m |\tilde{S}_j|^2 = O(1).$$

Hence, by (4.5) and (4.6), $R(G, d^G;\ \xi_{0,n}) = E\Phi(-\tau_0(\lambda, \psi)) + o(1)$.

REFERENCES

BROWN, L. D. and LOW, M. G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398.

BROWN, L. D. and ZHANG, C.-H. (1996). Coupling inequalities for some random design matrices and asymptotic equivalence of nonparametric regression and white noise. Preprint.

EFROMOVICH, S. and SAMAROV, A. (1996). Asymptotic equivalence of nonparametric regression and white noise has its limits. *Statist. Probab. Lett.* **28** 143–145.

NUSSBAUM, M. (1996). Asymptotic equivalence of density estimation and white noise. *Ann. Statist.* **24** 2399–2430.

ROBBINS, H. (1951). Asymptotically subminimax solutions of compound statistical decision problems. *Proc. Second Berkeley Symp. Math. Statist. Probab.* **1** 131–148. Univ. California Press, Berkeley.

DEPARTMENT OF STATISTICS
THE WHARTON SCHOOL
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PENNSYLVANIA 19104-6302
E-MAIL: lbrown@stat.wharton.upenn.edu

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
BUSCH CAMPUS
PISCATAWAY, NEW JERSEY 08855
E-MAIL: czhang@stat.rutgers.edu