ADAPTIVE SPLINE ESTIMATES FOR NONPARAMETRIC REGRESSION MODELS*

G. K. GOLUBEV AND M. NUSSBAUM

(*Translated by G. K. Golubev*)

I. Consider the problem of estimating a regression function $f(x) \in L_2(0,\ 1)$ under observations

$$y_i^n = f(x_i^n) + \sigma\xi_i^n, \qquad i = 1,\ \ldots,\ n, \quad x_i^n \in (0,\ 1), \tag{1}$$

where the ξ_i^n are independent standart Gaussian random variables, while the regressors x_i^n are deterministic and equally spaced, i.e., $x_i^n = (2i-1)/(2n)$. We suppose that the unknown function $f(\cdot)$ is sufficiently smooth; more precisely, it belongs to some Sobolev space $W_2^q = \{f \in L_2(0,\ 1)\colon\ D^q f \in L_2(0,\ 1)\}$. Here and henceforth $D^q f$ stands for the qth derivative of the function f.

One question arises immediately. Namely, which estimate of the regression is the best one? We should note that the question we deal with is completely different from finite-dimensional parametric estimation problems. For these problems the maximal likelihood estimate or the Bayes estimates are plausible candidates for the role of good ones. Usually these estimates are simultaneously asymptotically minimax for a sufficiently wide class of loss functions (see, e.g., [1]). In our problem a quadratic loss function

$$r_n(f,\ \hat{f}) = \frac{1}{n}\sum_{i=1}^{n}\left[f(x_i^n) - \hat{f}(x_i^n,\ Y_n)\right]^2$$

plays the key role. Here $\hat{f}(x,\ Y_n)$ is an estimate of the regression $f(x)$, which is based upon observations of the vector $Y_n = (y_1^n,\ \ldots,\ y_n^n)$ in (1).

Consider the subset $W_2^q(P,\ V) = \{f \in W_2^q\colon\ \|D^q f\|^2 \leqq P,\ \|f\|^2 \leqq V\}$ of W_2^q; here and henceforth $\|\cdot\|$ stands for the norm in $L_2(0,\ 1)$. Then an asymptotically minimax (a.m.) estimate on the set $W_2^q(P,\ V)$ is defined as an estimate such that the following relation holds:

$$\lim_{n\to\infty}\ \sup_{f\in W_2^q(P,V)} E_f r_n(f,\ f_n^*)\left[\inf_{\hat{f}}\ \sup_{f\in W_2^q(P,V)} E_f r_n(f,\ \hat{f})\right]^{-1} = 1. \tag{2}$$

Here $E_f(\cdot)$ stands for average with respect to the measure $P_f(\cdot)$ induced by observations y_i^n in (1) for fixed f, while the inf is taken over all estimates of the regression function.

The paper [2] has played a very significant role in the construction of a.m. estimates. It was shown there that for a quadratic loss function one can search for a.m. estimates among linear ones on ellipsoids in Hilbert space. A definitive solution of the problem is given in [3].

At first sight the linear a.m. estimates seem to be good candidates for nonparametric estimates corresponding to quadratic loss functions. However, it is not entirely true since these estimates depend explicitly on P and q [3]. Unfortunately, this pair is almost never known and it cannot be estimated via observations. Thus, the linear estimates prove to be not too attractive in practice. It should be said that an important step toward constructing good a.m. estimates was taken in [4]. That paper constructed an a.m. P-independent estimate for known q. More precisely it was shown that the estimate is an a.m. one provided a range of values of P/σ^2 is *a priori* known. The estimate is obtained on the basis of the GCV-criterion [5] and naturally it is nonlinear.

The dependence of the estimate on q, although not as critical as that on P, is nevertheless undesirable since q determines the order of the asymptotically minimax risk ($n^{-2q/(2q+1)}$).

*Received by the editors June 5, 1989.

The aim of this paper is to construct a.m. estimates that depend neither on P nor q. In what follows such estimates are called *adaptive.*

II. The space S_n^q of splines is intimately related to the Sobolev space W_2^2. Recall that the former is defined as the set of real functions s, subject to the following conditions with $n > q$:

i) s is a polynomial of degree $2q-1$ in each subinterval (x_i^n, x_{i+1}^n), $i = 1, \dots, n-1$;

ii) s is a polynomial of degree $q-1$ in $[0, x_1^n)$ and $(x_n^n, 1]$;

iii) $D^{2q-2}s$ is continuous in $[0, 1]$.

It is convenient to introduce the Demmler–Reinsch basis $\{\varphi_k^q(x)\}_1^n$, which is defined uniquely by the relations [6]

$$\frac{1}{n}\sum_{i=1}^{n}\varphi_k^q(x_i^n)\varphi_m^q(x_i^n) = \delta_{km},$$

$$\int_0^1 D^q\varphi_k^q(x)D^q\varphi_m^q(x) = \delta_{km}\lambda_k^n, \qquad 0 = \lambda_1^n = \cdots = \lambda_q^n < \lambda_{q+1} < \cdots < \lambda_n^n,$$

where δ_{km} is the Kronecker symbol.

We observe that if $f(x) = \sum_{k=1}^n c_k\varphi_k^q(x) \in S_n^q$, then $\|D^q f\|^2 = \sum_{k=1}^n \lambda_k^n c_k^2$. Therefore, the asymptotics behavior of the eigenvalues λ_k^n $(n \to \infty)$ plays an important part in what follows. This behavior is well known (see [4, Thm. 2.2]):

$$\lambda_{q+k}^n = \big(1 + o(1)\big)(\pi k)^{2q}, \qquad n \to \infty, \tag{3}$$

where $o(1)$ denotes a term that tends to zero uniformly in $k \in [k_{1n}, k_{2n}]$ for arbitrary sequences $k_{1n} \to \infty$ and $k_{2n} = o(n^{2/(2q+1)})$.

The construction of a.m. estimates relies essentially on a result pertaining to the asymptotically minimax property for linear estimates. Write

$$\{Y_n, \varphi_k^\beta\} = n^{-1}\sum_{i=1}^{n} y_i^n\varphi_k^\beta(x_i^n)$$

and define a linear estimate by

$$f_n(x, \beta, w) = \sum_{k=1}^{n} h_k(\beta, w)\langle Y_n, \varphi_k^\beta\rangle\, \varphi_k^\beta(x), \tag{4}$$

where

$$h_i(\beta, w) = \begin{cases} 1, & i \in [1, \beta], \\ 1 - [(i-\beta)]^\beta, & i \in [\beta + 1, w + \beta], \\ 0, & i > w + \beta. \end{cases} \tag{5}$$

We define also a functional $L_n^q[h, s] = \sum_{k=1}^\infty \{|1 - h_k|^2\langle s, \varphi_k^q\rangle^2 + \sigma^2 h_k^2/n\}$, $\{h_k\} \in l_2(1, \infty)$, $s \in S_n^q$.

LEMMA 1. *The following relations hold* ($as\ n \to \infty$) :

$$\begin{aligned} \inf_{\hat f}\ \sup_{f\in W_2^q(P,V)} E_f r_n(f, \hat f) &= \big(1 + o(1)\big) \sup_{f\in W_2^q(P,V)} E_f r_n(f, f^0) \\ &= \big(1 + o(1)\big) \sup_{f\in W_2^q(P,V)} L_n^q[h^0, s] = \big(1 + o(1)\big)\Delta(q)\big[Pn/(\pi^{2q}\sigma^2)\big]^{-2q/(2q+1)}, \end{aligned} \tag{6}$$

where $\Delta(q) = q/(q+1)[(2q+1)(q+1)/q]^{1/(2q+1)}$ *the linear estimate* $f_n^0(x, Y_n) = f_n(x, \beta, w_n^q)$, *the sequence* $h_i^0 = h_i(q, w_n^q)$, *and*

$$(7) \qquad w_n^q = \left[Pn/(\pi^{2q}\sigma^2)\right]^{1/(2q+1)} \left[(2q+1)(q+1)/q\right]^{1/(2q+1)}.$$

This result, which implies in particular that the estimate $f_n(x, \beta, w_n^q)$ is asymptotically minimax, can be easily proved by using (3) together with the lower bound in (2) (see also [3]).

Thus, it is plausible that the adaptive estimate should take the form (4). To determine it completely one has to form the pair (β, w) via observations y_i^n. The heuristic reasoning that allows one to do this is as follows.

One must try to choose the pair (β, w) to minimize the loss $r_n(f, f_n)$. Observe that

$$r_n(f, f_n) = \sum_{i=1}^{n} \left\{h_i^2(\beta, w)\langle Y_n, \varphi_i^\beta\rangle^2 - 2h_i(\beta, w)\langle Y_n, \varphi_i^\beta\rangle\langle f, \varphi_i^\beta\rangle\right\} + n^{-1}\sum_{i=1}^{n} f^2(x_i^n).$$

It is clear that one has to minimize only the first sum on the right-hand side with respect to β and w. We note, however, that the terms $\langle Y_n, \varphi_i^\beta\rangle\langle f, \varphi_i^\beta\rangle$ involved in the sum do depend on the unknown function $f(\cdot)$. The only way to overcome this unpleasant fact is to replace these terms by their unbiased estimates $\langle Y_n, \varphi_i^\beta\rangle^2 - \sigma^2/n$.

Thus, we arrive at the following construction of an adaptive estimate. Define the functional

$$l_n[Y_n, \beta, w] = \sum_{i=1}^{n} \left\{\left[h_i^2(\beta, w) - 2h_i(\beta, w)\right]\left\langle Y_n, \varphi_i^\beta \right\rangle + 2\sigma^2 h_i(\beta, w)/n\right\},$$

where $h_i(\beta, w)$ is given by (5), and find a pair of integers $(\overline{q}, \overline{w})$ such that

$$(\overline{q}, \overline{w}) = \operatorname*{arg\,min}_{q,\beta\in[1,n]} l_n[Y_n, \beta, w].$$

Then the adaptive estimate is $\bar{f}_n(x, Y_n) = f_n(x, \bar{\beta}, \overline{w})$.

The main result of the paper is as follows.

THEOREM 1. *The estimate* $\bar{f}_n(x, Y_n)$ *is asymptotically minimax for any* q *and* $P > 0$ *(see* (2)).

It should be noted that the estimate $\bar{f}_n(x, Y_n)$ depends explicitly on the noise variance σ^2. In many circumstances this dependence is apparently undesirable. To avoid it one can construct an estimate of σ^2 from the observations Y_n. This can be done in a variety of ways. Here we present only one of them. Put

$$\hat{\sigma}_n^2[Y_n, \beta, w] = n^{-1}\sum_{i=1}^{n} \left[y_i^n - f_n(x_i^n, \beta, w)\right]^2.$$

This quantity is an estimate for the noise variance σ^2, which involves the linear estimate $f_n(\cdot)$ for the function $f(\cdot)$ (see (4)). If we substitute $\widehat{\sigma_n^2}$ for σ^2 in the functional $l_n[Y_n, \beta, w]$ a new functional

$$l_n^*[Y_n, \beta, w] = \sum_{i=1}^{n} \left[h_i^2(\beta, w) - 2h_i(\beta, w)\right]\left\langle Y_n, \varphi_i^\beta\right\rangle^2 \times \left[1 + 2n^{-1}\sum_{m=1}^{n} h_m(\beta, w)\right] + 2n^{-2}\sum_{k=1}^{n} h_k(\beta, w)\sum_{i=1}^{n}(y_i^n)^2$$

is obtained.

We then proceed to construct the estimate $\bar{f}_n(\cdot)$ analogously. Define a pair of integers

$$(q^*, w^*) = \operatorname*{arg\,min}_{q,\ \beta\in[1,n^{1-\gamma}]} l_n^*[Y_n, \beta, w], \qquad \gamma > 0,$$

and the estimate $f_n^*(x, Y_n) = f_n(x, \beta^*, w^*)$.

THEOREM 2. *If $\gamma \in (0, \frac{1}{2})$, then the estimate $f_n^*(x, Y_n)$ is asymptotically minimax for any fixed $P > 0$ and q.*

III. The proofs of Theorems 1 and 2 rest on rather simple properties of monotone sequences. Denote by H_n the set of all sequences in $l_2[1, \infty)$ subject to the following constraints:

i) $h_1 = 1,\ h_k \in [0, 1]$,

ii) $h_{k+1} \leqq h_k$,

iii) $h_k = 0,\ k \geqq n$.

Let $\|\xi_{im}\|$, $i, m \in [1, \infty)$, be a matrix of standard Gaussian random variables such that for fixed m the random variables ξ_{im}, $i \in [1, \infty)$, are independent. Moreover, let there be given a random sequence $\{h_k\}_1^\infty$ and an integer-valued random variable ν such that $P\{\{h_k\}_1^\infty \in H_n\} = 1$ and $P\{\nu \in [1, n]\} = 1$.

These random quantities can depend arbitrarily on the Gaussian matrix $\|\xi_{im}\|$. Let there also be given a deterministic matrix $\|a_{im}\|$, $i, m \in [1, \infty)$, such that $\sum_{i=1}^\infty a_{im}^2 = 1$, $m \in [1, \infty)$.

LEMMA 2. *For any $r > 0$ the following inequalities hold:*

$$E\left|\sum_{k=1}^\infty h_k\left[\xi_{k\nu}^2 - 1\right]\right| \leqq C(r)\, n^r \left[E\sum_{k=1}^\infty h_k^2\right]^{1/2}, \tag{8}$$

$$E\left|\sum_{k=1}^\infty (1-h_k)\xi_{k\nu} a_{k\nu}\right| \leqq C(r)\, n^r F\left[E\sum_{k=1}^\infty (1-h_k)^2 a_{k\nu}^2\right], \tag{9}$$

where $C(r)$ is a constant and $F(x) = [x \log(x^{-1}+4)\log\log(x^{-1}+16)]^{1/2}$.

Proof. Define random variables

$$k_m = \max_{t \geqq 1}\left\{\left|\sum_{k=1}^t \left[\xi_{km}^2 - 1\right]\right| \Big/ F_1(t)\right\},$$

where $F_1(x) = [x\log\log(x+16)]^{1/2}$ satisfies that simple inequalities

$$\begin{aligned}\sum_{k=1}^n \left[F_1(k) - F_1(k-1)\right]^2 &\leqq \sqrt{\log\log(16)} + \sum_{k=2}^n \left[F_1(k) - F_1(k-1)\right]^2 \\ &\leqq C + \sum_{k=2}^n \left[F_1'(k-1)\right]^2 \leqq C + \left(\log\log(n-1) + 1\right)\sum_{k=1}^{n-1} k^{-1} \\ &\leqq C \log(n)\log\log(n).\end{aligned}$$

Here and henceforth C stands for a constant whose value is inessential. Then, by using summation by parts and the monotonicity of $F_1(t)$, we obtain

$$\begin{aligned}\left|\sum_{k=1}^\infty h_k\left[\xi_{k\nu}^2 - 1\right]\right| &= \left|\sum_{k=1}^\infty (h_k - h_{k+1})\sum_{s=1}^k [\xi_{s\nu}^2 - 1]\right| \\ &\leqq \max_{m\leqq n} k_m \sum_{k=1}^\infty (h_k - h_{k+1})F_1(k) = \max_{m \leqq n} k_m \sum_{k=1}^\infty h_k\left[F_1(k) - F_1(k-1)\right] \\ &\leqq \max_{m\leqq n} k_m \left[\sum_{k=1}^\infty h_k^2\right]^{1/2}\left[\sum_{k=1}^n \left[F_1(k) - F_1(k-1)\right]^2\right]^{1/2} \\ &\leqq C\max_{m\leqq n} k_m \left[\sum_{k=1}^\infty h_k^2\right]^{1/2}\left[\log(n)\log\log(n)\right]^{1/2}.\end{aligned} \tag{10}$$

A standard technique employed in the proof of the law of the iterated logarithm (see, e.g., [7]) can readily be used to show that $E|k_m|^R \leqq C(R)$ for any $R \geqq 1$. Therefore, by Jensen's inequality, we have for $R > 2$

$$E\left[\max_{m\leqq n} k_m\right]^2 \leqq E\left[\sum_{s=1}^{n} |k_s|^R\right]^{2/R} \leqq \left[E\sum_{s=1}^{n} |k_s|^R\right]^{2/R} \leqq C(R)n^{2/R}.$$

From this, (10) and the Cauchy-Bunyakovsky inequality we obtain (8).

The proof of (9) is a accomplished in a similar way. Introduce the following notation:

$$E_j^m = \sum_{k=j}^{\infty} a_{km}^2, \qquad F_0(x) = \left[x \log\log(x^{-1}+16)\right]^{1/2},$$

$$\eta = \max_{m\leqq n}\max_{t\geqq 1}\left|\sum_{s=t}^{\infty} a_{sm}\xi_{sm}\right| \Big/ F_0(E_t^m).$$

Again by summation by parts, we obtain ($h_1 = 1$)

$$\left|\sum_{k=1}^{\infty}(1-h_k)\xi_{k\nu}a_{k\nu}\right| = \left|\sum_{k=2}^{\infty}(h_{k-1}-h_k)\sum_{s=k}^{\infty}\xi_{s\nu}a_{s\nu}\right|$$

$$\leqq \eta\sum_{k=2}^{\infty}(h_{k-1}-h_k)F_0(E_k^\nu) = \eta\sum_{k=1}^{\infty}(1-h_k)\left[F_0(E_k^\nu) - F_0(E_{k+1}^\nu)\right]. \tag{11}$$

Introduce now a linear functional $\Psi[v] = \sum_{k=1}^{\infty} v_k[F_0(E_k^\nu) - F_0(E_{k+1}^\nu)]$; we find its maximum over the set

$$A_D = \left\{v_s\colon \sum_{s=1}^{\infty} v_s^2\alpha_{s\nu}^2 \leqq D,\ v_l \leqq v_{l+1},\ v_l \in [0,\ 1]\right\}, \qquad D \leqq 1.$$

Since A_D is convex, there exists an extremal sequence $\{v_s^*\}_1^\infty$ for which the functional attains its maximum.

Therefore, since $F_0(t)$ is a convex monotone function of $t \in [0, 1]$, we find that

$$v_k^* = \min\left\{1,\ \mu\left[F_0(E_k^\nu) - F_0(E_{k+1}^\nu)\right]a_{k\nu}^{-2}\right\},$$

where μ is a root of the equation

$$\sum_{k=1}^{\infty}\alpha_{k\nu}^2 \min\left\{1,\ \mu^2\left[F_0(E_k^\nu) - F_0(E_{k+1}^\nu)\right]^2 a_{k\nu}^{-4}\right\} = D.$$

Introduce an integer $N(\mu)$ as follows:

$$N(\mu) = \max\left\{k > 0\colon \mu\left[F_0(E_k^\nu) - F_0(E_{k+1}^\nu)\right]a_{k\nu}^{-2} \leqq 1\right\}.$$

Then

$$\begin{aligned}\Psi[v^*] &= \mu\sum_{k=1}^{N(\mu)}\left[F_0(E_k^\nu) - F_0(E_{k+1}^\nu)\right]^2 a_{k\nu}^{-2} + \sum_{k=N(\mu)+1}^{\infty}\left[F_0(E_k^\nu) - F_0(E_{k+1}^\nu)\right],\\ &\mu^2\sum_{k=1}^{N(\mu)}\left[F_0(E_k^\nu) - F_0(E_{k+1}^\nu)\right]^2 a_{k\nu}^{-2} + E_{N(\mu)+1}^\nu = D.\end{aligned} \tag{12}$$

It follows from the latter equality that $E^{\nu}_{N(\mu)+1} \leqq D$ and

$$\mu^2 \leqq D \Big/ \sum_{k=1}^{N(\mu)} \left[F_0(E^{\nu}_k) - F_0(E^{\nu}_{k+1})\right]^2 a_{k\nu}^{-2}.$$

Thus, we have

$$\Psi[v^*] \leqq \left(D \sum_{k=1}^{N(\mu)} \left[F_0(E^{\nu}_k) - F_0(E^{\nu}_{k+1})\right]^2 a_{k\nu}^{-2} \right)^{1/2} + F_0(D). \tag{13}$$

We note next that the definition of the number $N(\mu)$ immediately implies the inequality $E^{\nu}_{N(\mu)+1} \geqq C\mu^2$. Furthermore, by definition of $F_0(\cdot)$, the following inequalities hold:

$$\begin{aligned}
\sum_{k=1}^{N(\mu)} \left[F_0(E^{\nu}_k) - F_0(E^{\nu}_{k+1})\right]^2 a_{k\nu}^{-2} &\leqq C \log\log\left(1/E^{\nu}_{N(\mu)+1} + 16\right) \\
&\quad \times \sum_{s=1}^{N(\mu)} \alpha_{s\nu}^2 / E^{\nu}_k \leqq C \log\log\left(1/E^{\nu}_{N(\mu)+1} + 16\right) \sum_{s=1}^{N(\mu)} \left(1 - E^{\nu}_{k+1}/E^{\nu}_k\right) \\
&\leqq C \log\log\left(1/E^{\nu}_{N(\mu)+1} + 16\right) \left[1 - \sum_{s=1}^{N(\mu)} \log\left(E^{\nu}_{k+1}/E^{\nu}_k\right)\right] \\
&\leqq C \log\log\left(1/E^{\nu}_{N(\mu)+1} + 16\right) \log\left(1/E^{\nu}_{N(\mu)+1} + 4\right). \qquad (14)
\end{aligned}$$

Thus, we find from (13) that

$$D \leqq E^{\nu}_{N(\mu)+1} + C \log\log\left(1/E^{\nu}_{N(\mu)+1} + 16\right) \log\left(1/E^{\nu}_{N(\mu)+1} + 4\right) E^{\nu}_{N(\mu)+1}.$$

Hence, $E^{\nu}_{N(\mu)+1} \geqq CD[\log\log(1/D + 16)\log(1/D + 4)]^{-1}$ Therefore, we obtain from (14) the inequality $\Psi[v^*] \leqq CF(D)$. This means that the following inequality holds:

$$\sum_{k=1}^{\infty} (1 - h_k)\left[F_0(E^{\nu}_k) - F_0(E^{\nu}_{k+1})\right] \leqq CF\left[\sum_{s=1}^{\infty} (1 - h_s)^2 \alpha_{s\nu}^2\right].$$

Thus, since $F^p(x)$ is convex for $p < 2$ and $x \in [0, 1]$, it follows from (11) and the Hölder and Jensen inequalities that

$$E\left|\sum_{k=1}^{\infty} (1 - h_k)\xi_k a_{k\nu}\right| \leqq C\left[E\eta^3\right]^{1/3} F\left[E \sum_{k=1}^{\infty} (1 - h_k)^2 a_{k\nu}^2\right].$$

The rest of the calculations are similar to those already employed in the proof of (8). One should only observe that the distribution of the random variables $\sum_{s=t}^{\infty} a_{sm}\xi_{sm}$ coincides with that of $b(E^m_T)$, where $b(t)$, $t \geqq 0$, is the standard Wiener process.

Proof of Theorem 1. In view of Lemma 1 it suffices to establish an upper bound for the minimax risk of the estimate $f_n(x, Y_n)$:

$$R_n = \sup_{f \in W_2^q(P,V)} E_f r_n(f, \bar{f}_n).$$

Since $r_n(f, \bar{f}_n)$ depends only on the values of $f(x_i^n)$, $i = 1, \ldots, n$, the equality $E_f r_n(f, \bar{f}_n) = E_{s_n} r_n(s_n, \bar{f}_n)$ holds, where $s_n \in S_n^q$ is the interpolating spline through the points x_i^n. The extremal property $\|D^q s_n\| \leqq \|D^q f\|$ of the spline is well known. Moreover, it is clear that $\|s_n\|^2 \leqq \|f\|^2 + Cn^{-1}[\|f\|^2 + \|D^q f\|^2]$. Hence, from some $n > n_0(P, V)$ onward,

$$R_n \leqq \sup_{s \in S_n^q(P,V)} E_s r_n(s, \bar{f}_n), \tag{15}$$

where $S_n^q(P, V) = \{s \in S_n^q\colon \|D^q s\|^2 \leqq P,\ \|s\|^2 \leqq 2V\}$.

Now we use relations connecting $r_n(s, f_n(\cdot,\beta,w))$, $l_n[Y_n,\beta,w]$ and $L_n^\beta[h(\beta,w),s]$. For brevity we write $\zeta_l^\beta = n^{-1/2}\sum_{i=1}^n \xi_i\varphi_l^\beta(x_i^n)$. Then

$$r_n(s,\ f_n(\cdot,\ \beta,\ w)) = l_n[Y_n,\ \beta\ ,w] + n^{-1}\sum_{i=1}^n s^2(x_i^n) + \sum_{i=1}^\infty h_i(\beta,\ w)\Big[2\sigma s_i^\beta\big\langle s,\varphi_i^\beta\big\rangle n^{-1/2} + \sigma^2\big((\zeta_i^\beta)^2 - 1\big)/n\Big] \tag{16}$$

and

$$l_n[Y_n,\ \beta,\ w] = L_n^\beta\big[h(\beta,\ w),\ s\big] - n^{-1}\sum_{i=1}^n s^2(x_i^n) + \sum_{i=1}^\infty\Big[h_i^2(\beta,\ w) - 2h_i(\beta,\ w)\Big]\Big[2\sigma\zeta_i^\beta\big\langle s,\ \varphi_i^\beta\big\rangle n^{-1/2} + \sigma^2\big((\zeta_i^\beta)^2 - 1\big)/n\Big]. \tag{17}$$

We need an estimate of the last sums in (16), (17). By Lemma 2, for any $r > 0$ we have

$$E\Bigg|\sum_{i=1}^\infty h_i(\beta,\ w)\Big[2\sigma\zeta_i^\beta\big\langle s,\varphi_i^\beta\big\rangle n^{-1/2} + \sigma^2\big((\zeta_i^\beta)^2-1\big)/n\Big]\Bigg| \leqq C(r)\,n^{r-1/2}\Big[E_s L_n^\beta\big[h(\beta,\ w),\ s\big]\Big]^{1/2}, \tag{18}$$

$$E\Bigg|\sum_{i=1}^\infty\Big[h_i^2(\beta,\ w) - 2h_i(\beta,\ w)\Big]\Big[2\sigma\zeta_i^\beta\big\langle s,\ \varphi_i^\beta\big\rangle n^{-1/2} + \sigma^2\big((\zeta_i^\beta)^2-1\big)/n\Big]\Bigg| \leqq C(r)\,n^{r-1/2}\Big[E_s L_n^\beta\big[h(\beta,\ w),\ s\big]\Big]^{1/2}. \tag{19}$$

In deriving these inequalities, we used monotonicity of the function $F(\cdot)$ as well as the simple inequality $L_n^\beta[h(\beta,\ w),\ s] \geqq \sigma^2/n$. From (16) and (18) we arrive at the relations

$$\begin{aligned} Er_n(s,\ \bar f_n) &= El_n[Y_n,\ \bar\beta,\ \bar w] + n^{-1}\sum_{i=1}^n s^2(x_i^n) + C(r)\,n^{r-1/2}\Big[E_s L_n^{\bar\beta}\big[h(\bar\beta,\ \bar w),\ s\big]\Big]^{1/2} \\ &\leqq l_n\big[Y_n,\ q,\ w_n^q\big] + n^{-1}\sum_{i=1}^n s^2(x_i^n) + C(r)\,n^{r-1/2}\big[E_s L_n^{\bar\beta}[h(\bar\beta),\ \bar w,\ s]\big]^{1/2} \\ &\leqq L_n^q\big[h(q,\ w_n^q),\ s\big] + C(r)\,n^{r-1/2}\big[E_s L_n^{\bar\beta}[h(\bar\beta,\ \bar w),\ s]\big]^{1/2}. \end{aligned} \tag{20}$$

Now we estimate the value of $E_s L_n^{\bar\beta}[h(\bar\beta,\ \bar w),\ s]$. Again, by virtue of relations (16)–(19), we find

$$\begin{aligned} E_s L_n^{\bar\beta}[h(\bar\beta,\ \bar w),\ s] &= E\,l_n[Y_n,\ \bar\beta,\ \bar w] + n^{-1}\sum_{i=1}^n s^2(x_i^n) \\ &\quad + C(r)\,n^{r-1/2}\Big[E_s L_n^{\bar\beta}\big[h(\bar\beta,\ \bar w),\ s\big]\Big]^{1/2} \leqq l_n[Y_n,\ q,\ w_n^q] + n^{-1}\sum_{i=1}^n s^2(x_i^n) \\ &\quad + C(r)\,n^{r-1/2}\Big[E_s L_n^{\bar\beta}\big[h(\bar\beta,\ \bar w),\ s\big]\Big]^{1/2} \leqq L_n^q\big[h(q,\ w_n^q),\ s\big] \\ &\quad + C(r)\,n^{r-1/2}\Big[E_s L_n^{\bar\beta}\big[h(\bar\beta,\ \bar w),\ s\big]\Big]^{1/2}. \end{aligned} \tag{21}$$

This yields immediately

$$E_s L_n^{\bar\beta}\left[h(\bar\beta,\ \overline{w}),\ s\right] \leqq 2L_n^q\left[h(q,\ w_n^q),\ s\right] + C(r)n^{2r-1}.$$

Therefore, by inserting this inequality into (20), we arrive at the following relation:

$$Er_n(s,\ \bar f_n) \leqq L_n^q\left[h(q,\ w_n^q),\ s\right] + C(r)n^{r-1/2}\left[L_n^q\left[h(q,\ w_n^q),\ s\right]\right]^{1/2} + C(r)n^{2r-1}.$$

Thus, (see (15))

$$\begin{aligned} R_n \leqq & \sup_{s\in S_n^q(P,V)} L_n^q\left[h(q,\ w_n^q),\ s\right] \\ & + C(r)n^{r-1/2}\left[\sup_{s\in S_n^q(P,V)} L_n^q\left[h(q,\ w_n^q),\ s\right]\right]^{1/2} + C(r)\,n^{2r-1}. \end{aligned}$$

Hence, if we take r to be sufficiently small ($r < 1/(2q+1)$) and apply Lemma 1, we arrive at the inequality ($n \to \infty$)

$$R_n \leqq \left(1+o(1)\right) \sup_{s\in S_n^q(P,V)} L_n^q\left[h(q,\ w_n^q),\ s\right].$$

This relation just means that the estimate $f_n(\cdot)$ is asymptotically minimax (see (6)).

The proof of Theorem 2 goes through almost verbatim as that of Theorem 1. One should only observe the following relation:

$$\begin{aligned} l_n^*[Y_n,\ \beta,\ w] = & \ l_n[Y_n,\ \beta,\ w] + 2n^{-1}\sum_{i=1}^{\infty} h_i(\beta,\ w)\Bigg[r_n\left(f,\ f_n(\cdot,\ \beta,\ w)\right) \\ & + \sigma^2 n^{-1}\sum_{k=1}^{n}(\xi_k^2-1) + 2\sigma^2 n^{-1}\sum_{k=1}^{\infty} h_k(\beta,\ w) + 2\sigma n^{-1/2} \\ & \times \sum_{k=1}^{\infty}\left(1-h_k(\beta,\ w)\right)\left\langle f,\ \varphi_k^\beta\right\rangle\zeta_k^\beta + 2\sigma^2 n^{-1}\sum_{k=1}^{\infty} h_k(\beta,\ w)(\xi_k^\beta - 1)\Bigg]. \end{aligned} \tag{22}$$

The last pair of terms in the relation can be estimated by means of Lemma 2:

$$\begin{aligned} & E_s\left|n^{-1}\sum_{k=1}^{\infty}\left(1-h_k(\beta,\ w)\right)\left\langle f,\varphi_k^\beta\right\rangle\xi_k^\beta + 2\sigma^2 n^{-1}\sum_{k=1}^{\infty} h_k(\beta,\ w)\left(\left(\xi_k^\beta\right)^2-1\right)\right| \\ & \leqq C(r)\,n^{r-1/2}\left[E_s L_n^\beta[h(\beta,w),s]\right]^{1/2}, \end{aligned} \tag{23}$$

where r is an arbitrary positive number.

We also observe that $h_k(\beta,\ w) = 0$ for $k \geqq 3\sum_{s=1}^\infty h_s^2(\beta,\ w)$. This implies in particular that

$$\sigma^2 n^{-1}\sum_{k=1}^{\infty} h_k(\beta,\ w) \leqq 3\sigma^2 n^{-1}\sum_{k=1}^{\infty} h_k^2(\beta,\ w) \leqq 3L_n^\beta\left[h(\beta,\ w),\ s\right]$$

and

$$\sigma^2 n^{-1}\sum_{i=1}^{\infty} h_i(\beta,\ w)\sum_{k=1}^{\infty}\left(\xi_k^2-1\right) \leqq C n^{-1/2}L_n^\beta\left[h(\beta,\ w),\ s\right].$$

We note also that by virtue of the hypothesis the theorem,

$$n^{-1}\sum_{k=1}^{\infty} h_k(\beta,\ w) = o(1), \qquad n\to\infty,$$

uniformly in β, $w \in [1, n^{1-\gamma}]$.

These relations allow one to rewrite (20) in the following way:

$$E_s r_n(s,\ f_n^*) = \big(1+o(1)\big)\, L_n^q\Big[h\big(q,\ w_n^q\big),\ s\Big] + C(r)n^{r-1/2}\Big[E_s L_n^{\beta^*}\big[h(\beta^*,\ w^*),\ s\big]\Big]^{1/2} + o(1)\, E_s r_n\big(s,\ f_n^*\big) + o(1)\, E_s L_n^{\beta^*}\big[h(\beta^*,\ w^*),\ s\big].$$

The inequality (21) takes in turn the following form:

$$E_s L_n^{\beta^*}\big[h(\beta^*,\ w^*),\ s\big] \leqq \big(1+o(1)\big) L_n^q\Big[h\big(q,\ w_n^q\big),\ s\Big] + C(r)\, n^{r-1/2} \times \Big[E_s L_n^{\beta^*}\big[h(\beta^*,\ w^*),\ s\big]\Big]^{1/2} + o(1) E_s r_n(s,\ f_n^*) + o(1) E_s L_n^{\beta^*}\big[h(\beta^*,\ w^*),\ s\big].$$

Then, in the same way as in the proof of Theorem 1 one can easily obtain the required result.

REFERENCES

[1] I. A. IBRAGIMOV AND R. Z. KHASMINSKY, *Statistic Estimation:Asymptotic Theory*, Springer-Verlag, Berlin-New York, 1981.
[2] M. S. PINSKER, *Optimal filtering of square integrable signals with Gaussian noise*, Problems of Information Transmission, 16 (1980), pp. 52–68. (In Russian.)
[3] M. NUSSBAUM, *Spline smoothing in regression models and asymptotic efficiency in L_2*, Ann. Statist., 13 (1985), pp. 984–997.
[4] P. SPECKMAN, *Spline smoothing and optimal rates of convergence in nonparametric regression models*, Ann. Statist., 13 (1985), pp. 970–983.
[5] P. GRAVEN AND G. WAHBA, *Smoothing noisy data with spline functions*, Numerical Math., 31 (1979), pp. 377–403.
[6] A. DEMMLER AND C. REINSCH, *Oscillation matrices with spline smoothing*, Numerical Math., 24 (1975), pp. 357–382.
[7] V. V. PETROV, *Sums of Independent Rrandom Variables*, Springer-Verlag, Berlin–New York, 1975.

PRESERVATION OF TYPE UNDER MIXING*

S. JANKOVIC

Abstract. We give a short survey on the characterizations of probability distributions that satisfy the property that the sum of a random number of independent identically distributed (i.i.d.) random variables is of the same type as one random variable from that sum, together with some related new results.

Let X_1, X_2, ... be a sequence of i.i.d. random variables,

$$S_n = \sum_{j=1}^{n} X_j, \qquad P(X_1 < x) = F(x), \tag{1}$$

where positive integer-valued random variable ν_p, $p \in (0, 1)$, does not depend on X_1, X_2, ..., $P\{\nu_p = n\} = p_n$, $\sum_{n=1}^{\infty} p_n = 1$. The problems we deal with are connected with characterizations of distributions satisfing the following condition: for each $0 < p < 1$ there exist real

*Received by the editors January 4, 1990.