The Annals of Statistics
1993, Vol. 21, No. 1, 235–254

TESTING FOR ADDITIVITY OF A REGRESSION FUNCTION

BY DANIEL BARRY

University College, Cork

Observations y_{ij} are made at points (x_{1i}, x_{2j}) according to the model $y_{ij} = F(x_{1i}, x_{2j}) + e_{ij}$, where the e_{ij} are independent normals with constant variance.

In order to test that $F(x_1, x_2)$ is an additive function of x_1 and x_2, a likelihood ratio test is constructed comparing

$$F(x_1, x_2) = \mu + Z_1(x_1) + Z_2(x_2)$$

with

$$F(x_1, x_2) = \mu + Z_1(x_1) + Z_2(x_2) + Z(x_1, x_2),$$

where Z_1, Z_2 are Brownian motions and Z is a Brownian sheet. The ratio of Brownian sheet variance to error variance α is chosen by maximum likelihood and the likelihood ratio test statistic W of H_0: $\alpha = 0$ used to test for departures from additivity.

The asymptotic null distribution of W is derived, and its finite sample size behaviour is compared with two standard tests in a simulation study. The W test performs well on the five alternatives considered.

1. Introduction. Let $(y_i, \mathbf{x}_i)$, $i = 1, 2, \ldots, N$ satisfy

$$y_i = F(\mathbf{x}_i) + e_i,$$

where $\mathbf{x}_i \in R^d$ for each i, F is a fixed but unknown regression function, and the errors $\{e_i\}$ are uncorrelated with mean zero and variance σ^2. The problem of estimating the function F has a huge literature associated with it. Methods such as kernel and nearest neighbour, found to be useful in one dimension, have been extended to provide estimates of functions of many variables. [See Prakasa Rao (1983) for a review.] Projection pursuit regression approximates F by a sum of univariate functions, each function depending on a particular linear combination of the elements of $\mathbf{x}$. [See Friedman and Stuetzle (1981).] The additive models of Buja, Hastie and Tibshirani (1989) approximate F by a sum of d univariate functions, one for each dimension of $\mathbf{x}$.

Interaction spline models were introduced in Barry (1983) and were further developed by Barry (1986, 1988), Wahba (1986) and Chen (1987, 1991). Borrowing ideas from the theory of analysis of variance, an interaction spline model writes F as a constant term plus a sum of d functions of one variable (the "main effects") plus a sum of $d(d-1)/2$ functions of two variables (the

Received January 1990; revised February 1992.
AMS 1991 *subject classifications*. Primary 62G10; secondary 62E20, 62M10.
Key words and phrases. Nonparametric regression, Brownian sheet, nonstandard likelihood asymptotics, tests for additivity.

"two factor interactions") and so on. For $d = 2$ the decomposition is given by

$$F(x_1, x_2) = \mu + a(x_1) + b(x_2) + c(x_1, x_2),$$

where

$$\mu = \int\int F(u, v)\, du\, dv,$$

$$(1.1) \qquad a(x_1) = \int F(x_1, v)\, dv - \mu,$$

$$b(x_2) = \int F(u, x_2)\, du - \mu,$$

$$c(x_1, x_2) = F(x_1, x_2) - \alpha(x_1) - \beta(x_2) - \mu.$$

F is estimated by the function $\hat{F}$ which minimizes the sum of squared residuals plus a weighted sum of penalty terms quantifying the "roughness" of each of the functions used in the ANOVA decomposition.

In d dimensions the ANOVA decomposition involves 2^d different functions and to estimate them all would require a large amount of data and would result in a fitted function which would not be easy to interpret. Similar problems arise in multifactor analysis of variance and stepwise model building procedures are available there to enable parsimonious models to be fitted when appropriate. This paper is part of an endeavour to provide a similar technology in the case of interaction splines. Chen (1987) considers fitting periodic interaction splines and proposes a stepwise model building procedure using the generalized cross validation function of Craven and Wahba (1979). In this paper we consider an alternative approach which exploits the connection between splines and Bayes estimates described in Wahba (1978). We make a detailed examination of the two-dimensional grid case. The methodology may be applied to more complicated multi-dimensional problems using the numerical algorithms of Gu, Bates, Chen and Wahba (1988).

Suppose that our observations are made at points on the grid

$$\{(x_{ij}, x_{2i}) \colon 1 \le i \le N_1, 1 \le j \le N_2\}$$

and that

$$y_{ij} = F(x_{ij}, x_{2j}) + e_{ij},$$

where the errors $\{e_{ij}\}$ are uncorrelated with mean zero and variance v. Barry (1986) makes use of the ANOVA decomposition of (1.1) to propose a probability model for $\{y_{ij}\}$ which includes the assumptions that the quantities

$$F_{ij} = F(x_{1i+1}, x_{2j+1}) + F(x_{1i}, x_{2j}) - F(x_{1i+1}, x_{2j}) - F(x_{1i}, x_{2j+1})$$

are independent and normally distributed with mean zero and variance $v_{12}(x_{1i+1} - x_{1i})(x_{2j+1} - x_{2j})$ and that the errors $\{e_{ij}\}$ are iid $N(0, v)$. Clearly $v_{12} = 0$ forces F to be additive. The test for additivity we propose is the likelihood ratio test of the hypothesis H_0: $v_{12} = 0$ based on the observables

$$w_{ij} = y_{i+1j+1} + y_{ij} - y_{i+1j} - y_{ij+1},$$

whose distribution depends only on v_{12} and v.

A similar idea has been proposed by Barry and Hartigan (1990) to test the hypothesis of constant regression function given a sample of $\{(y_i, x_i): 1 \leq i \leq n\}$, and by Yaganimoto and Yaganimoto (1987) to test the adequacy of the fit of a simple linear regression model. Cox and Koh (1989) and Cox, Koh, Wahba and Yandell (1988) introduce the theory of locally most powerful tests in a more general setting.

The paper is organized as follows. The test statistic W is derived in Section 2. In Section 3 we derive the asymptotic distribution of the test statistic under the null hypothesis that F is additive for observations made on a uniform grid. Section 4 contains a simulation study comparing the power of the proposed test with that of two well-known tests for additivity, namely, Tukey's one degree-of-freedom test and a test proposed by Johnson and Graybill (1972). We conclude that the W test is superior to that of Johnson and Graybill and is more reliable than Tukey's test. Proofs of theorems from Section 3 are given in Section 5.

2. The test statistic. Consider the grid of $N = N_1 N_2$ points $\{(x_{1i}, x_{2j}): 1 \leq i \leq N_1, 1 \leq j \leq N_2\}$ with

$$0 \leq x_{11} < x_{12} < \cdots < x_{1N_1} \leq 1,$$

$$0 \leq x_{21} < x_{22} < \cdots < x_{2N_2} \leq 1.$$

Suppose we observe

$$y_{ij} = F(x_{1i}, x_{2j}) + e_{ij},$$

where $\{e_{ij}\}$ are random errors and F: $[0,1] \times [0,1] \to R$ is an unknown regression function. We wish to test the hypothesis that F is an additive function of x_1 and x_2.

Motivated by a decomposition often used in two-way analysis of variance we can write

$$F(x_1, x_2) = \mu + a(x_1) + b(x_2) + c(x_1, x_2),$$

where μ, a, b and c are defined in (1.1).

A prior for F is constructed by putting independent priors on μ, a, b and c as follows:

(i) $\mu \sim N(0, v_0)$.

(ii) $a(x_1) \sim Z_1(x_1) - \int_0^1 Z_1(u)\,du$, where Z_1 is a Brownian motion with variance v_1.

(iii) $b(x_2) \sim Z_2(x_2) - \int_0^1 Z_2(u)\,du$, where Z_2 is a Brownian motion with variance v_2.

(iv) $$c(x_1, x_2) \sim Z(x_1, x_2) - \int_0^1 Z(x_1, u)\,du - \int_0^1 Z(u, x_2)\,du + \int_0^1 \int_0^1 Z(u, v)\,du\,dv,$$

where Z is a Brownian sheet with variance v_{12}.

To complete the specification of the probability model for the data we assume

(v) given F, the data $\{y_{ij}\}$ are independent with

$$y_{ij} \sim N\big(F(x_{1i}, x_{2j}), v\big) \qquad 1 \le i \le N_1, 1 \le j \le N_2.$$

This model was introduced by Barry (1986) in the context of estimating the function F.

Let $n_1 = N_1 - 1$ and $n_2 = N_2 - 1$. For $i = 1, 2, \dots, n_1$ and $j = 1, 2, \dots, n_2$, define

$$w_{ij} = y_{i+1j+1} + y_{ij} - y_{i+1j} - y_{ij+1}.$$

Then under the probability model specified in (i)–(v),

$$\mathbf{w} = (w_{11}, w_{21}, \dots, w_{n_11}, w_{21}, \dots, w_{n_1n_2})$$

has a multivariate normal distribution with mean zero and covariance matrix $v\Sigma$, where

$$\Sigma = \alpha D_1 \times D_2 + E_1 \times E_2,$$

where

$$\alpha = v_{12}/v,$$

$$D_1 = \text{diag}\{x_{12} - x_{11}, x_{13} - x_{12}, \dots, x_{1N_1} - x_{1n_1}\},$$

E_1 is an $n_1 \times n_1$ tridiagonal matrix with all diagonal entries equal to 2 and off diagonal entries equal to -1, D_2 and E_2 are like D_1 and E_1

and $A \times B$ denotes the Kronecker product of A and B [see Bellman (1970)].

Let $L(\alpha, v)$ be the log likelihood for α and v based on $\mathbf{w}$. Let $\hat{v}(\alpha)$ be the maximum likelihood estimate of v given α. Define

$$Q(\alpha) = 2L(\alpha, \hat{v}(\alpha))$$

and let $\hat{\alpha}$ be the value for which $Q(\alpha)$ is a maximum. Clearly

$$\alpha = 0 \quad \Rightarrow \quad v_{12} = 0$$
$$\Rightarrow \quad F \text{ is additive}$$

and so to test the hypothesis that F is additive we propose the test statistic

$$W = Q(\hat{\alpha}) - Q(0)$$

and reject additivity for large values of W.

3. The uniform grid case. We now specialize to the case where

$$x_{1i+1} - x_{1i} = \frac{1}{N_1}, \qquad i = 1, 2, \dots, n_1,$$

$$x_{2j+1} - x_{2j} = \frac{1}{N_2}, \qquad j = 1, 2, \dots, n_2.$$

In this case

$$\Sigma = \frac{\alpha}{N_1 N_2} I + E_1 \times E_2$$

and hence

$$\begin{aligned} 2L(\alpha, v) &= -n_1 n_2 \log v - \log|\Sigma| - \frac{1}{v}\mathbf{w}'\Sigma^{-1}\mathbf{w} \\ \Rightarrow \quad \hat{v}(\alpha) &= \frac{\mathbf{w}'\Sigma^{-1}\mathbf{w}}{n_1 n_2} \\ \Rightarrow \quad Q(\alpha) &= 2L(\alpha, \hat{v}(\alpha)) \\ &= n_1 n_2 \log(n_1 n_2) - n_1 n_2 \log(\mathbf{w}'\Sigma^{-1}\mathbf{w}) \\ &\quad - \log|\Sigma| - n_1 n_2. \end{aligned}$$

It can be checked that E_1 has eigenvalues

$$\lambda_{1r} = 2\left(1 - \cos\left(\frac{\pi r}{N_1}\right)\right), \qquad r = 1, 2, \dots, n_1,$$

with corresponding unit eigenvectors

$$\mathbf{v}_r = (v_{ri}),$$

where

$$v_{ri} = D_r \sin\left(\frac{\pi r i}{N_1}\right), \qquad i = 1, 2, \dots, n_1,$$

where D_r is a normalizing constant. Similar results are true for E_2.

Hence

$$\log|\Sigma| = \sum\sum \log\left[\frac{\alpha}{N_1 N_2} + \lambda_{1r}\lambda_{2s}\right]$$

and

$$\mathbf{w}'\Sigma^{-1}\mathbf{w} = \sum\sum \frac{U_{rs}^2}{(\alpha/N_1 N_2) + \lambda_{1r}\lambda_{2s}},$$

where

$$U_{rs} = \sum\sum v_{ri} v_{sj} w_{ij}.$$

When $v_{12} = 0$ it can be shown that the quantitites $Z_{rs} = U_{rs}/(v\lambda_{1r}\lambda_{2s})^{1/2}$ are iid $N(0, 1)$. Hence we write

$$Q(\alpha) - Q(0) = -n_1 n_2 \log\left\{ \frac{\sum\sum \lambda_{1r}\lambda_{2s} Z_{rs}^2 / ((\alpha/N_1N_2) + \lambda_{1r}\lambda_{2s})}{\sum\sum Z_{rs}^2} \right\}$$

$$+ \sum\sum \log\left\{ \frac{\lambda_{1r}\lambda_{2s}}{\alpha/N_1N_2 + \lambda_{1r}\lambda_{2s}} \right\}.$$

For later convenience we write in terms of $\beta = N_1 N_2 \alpha$ to get

$$M(\beta) \equiv Q\left(\frac{\beta}{N_1 N_2} \right) - Q(0)$$

$$= -n_1 n_2 \log\left\{ \frac{\sum\sum N_1^2 N_2^2 \lambda_{1r}\lambda_{2s} Z_{rs}^2 / (\beta + N_1^2 N_2^2 \lambda_{1r}\lambda_{2s})}{\sum\sum Z_{rs}^2} \right\}$$

$$+ \sum\sum \log\left\{ \frac{N_1^2 N_2^2 \lambda_{1r}\lambda_{2s}}{\beta + N_1^2 N_2^2 \lambda_{1r}\lambda_{2s}} \right\}.$$

The test statistic W is given by $W = M(\hat{\beta})$, where $\hat{\beta}$ is the value for which $M(\beta)$ is a maximum. We now examine the large sample behaviour of $M(\hat{\beta})$. For ease of presentation we restrict attention to the case $N_1 = N_2 = N$. The asymptotic results we prove as $N \to \infty$ can be shown to be true also as $N_1 \to \infty$, $N_2 \to \infty$.

We have the following theorem.

THEOREM 1. *Let* $\{z_{rs}\colon 1 \le r < \infty,\ 1 \le s < \infty\}$ *be a collection of iid* $N(0,1)$ *random variables. Define*

$$M_0(\beta) = \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \frac{\beta z_{rs}^2}{\beta + \pi^4 r^2 s^2} + \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \log\left\{ \frac{\pi^4 r^2 s^2}{\pi^4 r^2 s^2 + \beta} \right\}.$$

Let $\tilde{\beta}$ *be the value for which* $M_0(\beta)$ *is a maximum. Then*

$$W - M_0(\tilde{\beta}) \to_P 0$$

as $N \to \infty$.

The proof is given in Section 5 and involves 3 steps. We show that:

(a) $\hat{\beta}$ is bounded with probability 1.
(b) $\tilde{\beta}$ is bounded with probability 1.
(c) $M(\beta)$ converges uniformly to $M_0(\beta)$ over any finite interval $[0, \beta_0]$.

These steps are covered by the following theorems whose proofs are in Section 5.

THEOREM 2. *Let $\hat{\beta}$ be the value which maximizes $M(\beta)$. Then*

$$\limsup_{N\to\infty} P\{\hat{\beta} \geq \beta_0\} \to 0 \quad \text{as } \beta_0 \to \infty.$$

THEOREM 3. *Let $M_0(\beta)$ be as in Theorem* 1. *Let $\tilde{\beta}$ be the value for which $M_0(\beta)$ is a maximum. Then*

$$P\{\tilde{\beta} \geq \beta_0\} \to 0 \quad \text{as } \beta_0 \to \infty.$$

THEOREM 4. *For any fixed $\beta_0 < \infty$,*

$$\sup_{0\leq\beta\leq\beta_0} |M(\beta) - M_0(\beta)| \to_P 0 \quad \text{as } N \to \infty.$$

Table 1 gives estimates of the 10, 5 and 1 percentiles of the null distribution of W for various sample sizes. All the estimate are based on 10,000 repetitions. For each repetition the value of $\hat{\beta}$ was obtained by global search over a

TABLE 1
The null distribution of W. Showing, for various sample sizes, estimates of the 10, 5 *and* 1 *percentiles and the percentage of zero values based on* 10,000 *repetitions*

	n_2				
n_1	5	10	20	40	100
			(a) The 10 percentiles		
5	1.43				
10	1.17	1.22			
20	1.16	1.12	1.08		
40	1.09	1.08	1.07	1.13	
100	1.09	1.11	1.04	1.18	1.11
			(b) The 5 percentiles		
5	2.52				
10	2.22	2.13			
20	2.07	2.08	2.01		
40	1.95	2.02	1.99	1.99	
100	2.03	1.98	2.08	2.08	2.06
			(c) The 1 percentiles		
5	5.25				
10	4.84	4.62			
20	4.67	4.37	4.72		
40	4.27	4.59	4.28	4.25	
100	4.46	4.33	4.16	4.41	4.52
			(d) The percentage of zero values		
5	60.2				
10	62.0	61.9			
20	62.1	62.7	63.2		
40	63.0	63.7	64.0	64.7	
100	62.6	63.4	65.5	66.6	71.7

wide grid of values followed by a more concentrated localized search once the neighbourhood of the maximum had been located. The quantities seem to have stabilized and the asymptotic null distribution been attained at about $n_1 = n_2 = 20$. The null distribution of W has a point mass of about 0.65 at zero. Thus any positive value of W produces a p value less than 0.35.

In general, the grid points will not be equally spaced; in this case the log likelihood has the same form except that the eigenvalues $\{\lambda_{1r}\}$ and $\{\lambda_{2s}\}$ depend on the grid spacings. Similar asymptotic calculations should be possible in this case also, under suitable conditions on the grid spacings.

4. Simulation study. In this section we report the results of a simulation study comparing the power of three tests for additivity.

4.1. *The W test* (W). This test is described in Section 2 and uses critical values obtained in the simulation experiment of Section 3.

4.2. *Tukey's one degree-of-freedom test* (T). This test was introduced by Tukey (1949) and is based on the test statistic

$$T = \sqrt{(N_1N_2 - N_1 - N_2)R/(R_0 - R)},$$

where

$$R_0 = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot}\right)^2,$$

$$R = \frac{\left[\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot}\right)\left(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}\right)\left(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot}\right)\right]^2}{\sum_{i=1}^{N_1}\left(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}\right)^2\sum_{j=1}^{N_2}\left(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot}\right)^2},$$

$$\bar{y}_{i\cdot} = \frac{1}{N_2}\sum_{j=1}^{N_2} y_{ij}, \qquad y_{\cdot j} = \frac{1}{N_1}\sum_{i=1}^{N_1} y_{ij}, \qquad \bar{y}_{\cdot\cdot} = \frac{1}{N}\sum\sum y_{ij}.$$

T has a t_M distribution with $M = N_1N_2 - N_1 - N_2$ if F is additive and additivity is rejected for large values of T.

4.3. *Johnson and Graybill test* (JG). Let B be the $N_1 \times N_2$ matrix with elements

$$y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot}.$$

Let $C = B^TB$ and let

$$\text{JG} = \frac{\text{Maximum eigenvalue of } C}{\text{Trace}(C)}.$$

Reject additivity for large values of JG. This test statistic was derived in Johnson and Graybill (1972) where critical values are also given.

The following seven test functions were used:

(a) $$F(x_1, x_2) = x_1 x_2$$

(b) $$F(x_1, x_2) = \frac{e^{5(x_1+x_2)} - 1}{e^{5(x_1+x_2)} + 1}.$$

(c) $$F(x_1, x_2) = (1 + \sin(2\pi(x_1 + x_2)))/2.$$

(d) $$F(x_1, x_2) = 64(x_1 x_2)^3(1 - x_1 x_2)^3.$$

(e) $$F(x_1, x_2) = G(x_1)G(x_2)/36,$$

where
$$G(x) = \begin{cases} 15x & 0 \le x \le 0.2, \\ 5 - 10x & 0.2 \le x \le 0.4, \\ -9 + 25x & 0.4 \le x \le 0.6, \\ 18 - 20x & 0.6 \le x \le 0.8, \\ -2 + 5x & 0.8 \le x \le 1. \end{cases}$$

(f) $$F(x_1, x_2) = \begin{cases} 1, & \text{if } x_1 > 0.5 \text{ and } x_2 > 0.5, \\ 0, & \text{otherwise.} \end{cases}$$

(g) $$F(x_1, x_2) = (x_1 + x_2)/2 + 1 \text{ outlier}.$$

In the simulation study the position of the outlier was chosen at random; the value of F at the outlier position was $(x_1 + x_2)/2 + 3$.

Six values for (N_1, N_2) were used: $(N_1, N_2) = (5,5), (5,10), (5,20), (10,10), (10,20)$ and $(20,20)$. Three values for SD, the standard deviation of the error distribution were used: 0.1, 0.5 and 1.0.

For each combination of F, (N_1, N_2) and SD, 1000 datasets were generated by setting

$$x_{1i} = \frac{2i - 1}{N_1},$$

$$x_{2j} = \frac{2j - 1}{N_2},$$

$$y_{ij} = F(x_{1i}, x_{2j}) + e_{ij},$$

for $1 \le i \le N_1$, $1 \le j \le N_2$, where the $\{e_{ij}\}$ are iid $N(0, \mathrm{SD}^2)$. The proportion of results significant at the 1, 5 and 10% levels was found for each of the three test procedures. Table 2 gives a sample of the results. The complete set of results are available upon request from the author.

The standard errors for differences between two percentages in the table never exceed 2%. (Note that the test statistics were each computed on the same data sample in order to reduce the variance of the difference between the percentages exceeding critical values.)

The Johnson and Graybill test performed poorly relative to the W test. It had uniformly lower power and was highly sensitive to the presence of outliers.

TABLE 2

Showing, for various values of N_1, N_2 and SD and various choices of F, the percentage of 1000 iterations for which the 5% critical value was exceeded using the W test (W), Tukey's test (T) and the Johnson–Graybill test (JG)

	$(N_1, N_2) = (5, 5)$			$(N_1, N_2) = (5, 20)$			$(N_1, N_2) = (20, 20)$		
SD	**W**	**T**	**JG**	**W**	**T**	**JG**	**W**	**T**	**JG**
				(a) $F(x_1, x_2) = x_1 x_2$					
0.1	98.4	99.0	30.5	100.0	100.0	99.8	100.0	100.0	100.0
0.5	11.3	7.1	2.0	36.2	14.6	5.4	89.8	64.3	8.2
1.0	7.3	3.6	2.6	10.5	5.7	6.0	34.0	8.8	6.4
				(b) $F(x_1, x_2) = \exp(5(x_1 + x_2))/(1 + \exp(5(x_1 + x_2))) - 1$					
0.1	65.6	57.1	11.3	100.0	100.0	89.6	100.0	100.0	100.0
0.5	8.3	5.4	3.1	18.2	5.1	5.3	68.9	17.7	8.7
1.0	5.9	4.5	2.4	8.0	5.0	5.2	21.0	6.0	5.6
				(c) $F(x_1, x_2) = 0.5(1 + \sin(2\pi(x_1 + x_2)))$					
0.1	100.0	4.7	0.0	100.0	5.3	79.8	100.0	6.3	100.0
0.5	30.5	5.0	1.5	99.8	4.5	18.1	100.0	4.9	99.9
1.0	9.9	4.6	2.6	59.5	5.0	7.0	99.9	5.0	24.5
				(d) $F(x_1, x_2) = 64(x_1 x_2)^3(1 - x_1 x_2)^3$					
0.1	18.4	63.7	6.2	66.6	100.0	51.5	100.0	100.0	100.0
0.5	6.9	4.8	2.0	7.4	7.9	5.3	11.5	23.2	5.9
1.0	5.7	5.9	3.3	6.3	4.0	5.4	5.4	6.2	6.4
				(e) $F(x_1, x_2)$ = product of sawtooths					
0.1	18.1	65.7	6.1	72.0	100.0	50.7	100.0	100.0	100.0
0.5	5.4	6.9	2.7	6.9	8.7	5.2	12.3	24.3	6.0
1.0	4.3	4.0	1.7	5.8	4.9	4.4	8.0	6.0	6.4
				(f) $F(x_1, x_2) = 1$ if $x_1 > 0.5$ and $x_2 > 0.5$; $F(x_1, x_2) = 0$ otherwise					
0.1	100.0	100.0	99.9	100.0	100.0	100.0	100.0	100.0	100.0
0.5	47.0	29.6	12.2	98.8	86.2	47.1	100.0	100.0	99.8
1.0	22.1	11.2	7.3	56.7	18.9	14.0	99.5	85.8	30.6
				(g) $F(x_1, x_2) = (x_1 + x_2)/2 + 1$ outlier					
0.1	16.7	99.4	100.0	12.7	69.1	100.0	12.7	22.6	100.0
0.5	13.8	46.0	42.9	10.3	25.5	56.4	10.8	11.4	44.4
1.0	9.9	13.8	9.5	12.2	12.9	14.5	11.1	9.8	13.5

The comparison between the W test and Tukey's test is less clearcut. The W test is marginally better for functions (a), (b) and (f) with Tukey's test better for (d) and (e). The greatest discrepancy occurs for (c), where the T test performs very badly. The poor performance of Tukey's test here is no surprise since $\int_0^1 F(x,y)\,dx = \int_0^1 F(x,y)\,dy = \frac{1}{2}$ and the power of Tukey's test is known to be low in the absence of row and column main effects [Hegemann and Johnson (1976)]. The results for (g) indicate that T is more sensitive to outliers than is W. We conclude that the W test, while not being uniformly more powerful than Tukey's test, is more reliable principally because it does not exhibit the disastrously low power of Tukey's test when row and column main effects are absent.

5. Proofs for Section 3. We shall make use of the following notation:

$$n = N - 1,$$

$$\lambda_r = 2\left(1 - \cos\frac{\pi r}{N}\right), \qquad r = 1, 2, \dots, n,$$

$$a_{rs} = N^4 \lambda_r \lambda_s,$$

$$b_{rs}(\beta) = (\beta + a_{rs})^{-1}.$$

When the range of summation is from 1 to n it will be suppressed for ease of notation. We will use GR to denote Gradsteyn and Ryzhik (1980).

PROOF OF THEOREM 2. We show separately that

$$\text{(a)} \qquad \limsup_{N\to\infty} P\{\hat{\beta} > 6N^4\} = 0$$

and

$$\text{(b)} \qquad \limsup_{N\to\infty} P\{\beta_0 \le \hat{\beta} \le 6N^4\} \to 0 \quad \text{as } \beta_0 \to \infty.$$

PROOF OF (a). Since $M(0) = 0$ it suffices to show that

$$\limsup_{N\to\infty} P\left\{\sup_{\beta > 6N^4} M(\beta) > 0\right\} = 0.$$

Since

$$b_{rs}(\beta) \le 1/\beta$$

and

$$\beta b_{rs}(\beta) \ge \beta b_{nn}(\beta) \ge 6N^4 b_{nn}(6N^4)$$

we have that

$$\sup_{\beta > 6N^4} \frac{M(\beta)}{n^2} \leq -\log\{6N^4 b_{nn}(6N^4)\}$$

$$- \log\left\{ \frac{\Sigma\Sigma a_{rs} Z_{rs}^2}{\Sigma\Sigma Z_{rs}^2} \right\} + \frac{1}{n^2} \sum\sum \log(a_{rs})$$

$$= -\log\{6N^4 b_{nn}(6N^4)\}$$

$$- \log\left\{ \frac{\Sigma\Sigma \lambda_r \lambda_s Z_{rs}^2}{\Sigma\Sigma Z_{rs}^2} \right\} + \frac{1}{n^2} \sum\sum \log(\lambda_r \lambda_s).$$

Since

(i) $$\lambda_n \to 4,$$

(ii) $$\frac{1}{n} \sum_{r=1}^{n} \log(\lambda_r) \to \frac{1}{\pi} \int_0^{\pi} \log(2 - 2\cos x)\, dx = 0,$$

(iii) $$\frac{\Sigma\Sigma \lambda_r \lambda_s Z_{rs}^2}{\Sigma\Sigma Z_{rs}^2} \to_P 4$$

because

$$\sum_{r=1}^{n} \cos\left(\frac{\pi r}{N} \right) = 0$$

we have that

$$P\left\{ \sup_{\beta > 6N^4} \frac{M(\beta)}{n^2} \leq \log\left(\frac{22}{24} \right) \right\} \to 1 \quad \text{as } N \to \infty$$

which proves (a). □

Proof of (b). It suffices to show that

$$\limsup P\left\{ \sup_{\beta_0 \leq \beta \leq 6N^4} M'(\beta) > 0 \right\} \to 0 \quad \text{as } \beta_0 \to \infty.$$

Since

$$M'(\beta) = \frac{n^2 \Sigma\Sigma a_{rs} b_{rs}^2(\beta) Z_{rs}^2}{\Sigma\Sigma a_{rs} b_{rs}(\beta) Z_{rs}^2} - \sum\sum b_{rs}(\beta),$$

we have that

$$M'(\beta) > 0 \quad \Leftrightarrow \quad \sum\sum c_{rs}(\beta) Z_{rs}^2 > 0,$$

where

$$c_{rs}(\beta) = a_{rs} b_{rs}^2(\beta) - a_{rs} b_{rs}(\beta) \left[\frac{1}{n^2} \sum\sum b_{rs}(\beta) \right].$$

Hence

$$P\left\{\sup_{\beta_0 \le \beta \le 6N^2} M'(\beta) > 0\right\} \le \sum_{\beta=\beta_0}^{6N^4} P\left\{\sum\sum d_{rs}(\beta) Z_{rs}^2 > 0\right\},$$

where $d_{rs}(\beta) = \sup_{\beta \le h \le \beta+1} c_{rs}(h)$.

We will prove the following lemma later.

Lemma 1. *There exist positive constants B_1, B_2 and N_0 such that*

(a) $$\sum\sum d_{rs}(\beta) < -B_1 \beta^{-1/2}$$

and

(b) $$\sum\sum d_{rs}^2(\beta) < B_2 \beta^{-3/2}$$

for $N_0 \le \beta \le 6N^4$.

By Chebyshev's inequality we have that for any even integer m

$$P\left\{\sum\sum d_{rs}(\beta) Z_{rs}^2 > 0\right\} \le \frac{E\left(\Sigma\Sigma d_{rs}(\beta)\left(Z_{rs}^2 - 1\right)\right)^m}{\left(\Sigma\Sigma d_{rs}(\beta)\right)^m}.$$

Whittle (1960) shows that

$$E\left(\sum\sum d_{rs}(\beta)\left(Z_{rs}^2 - 1\right)\right)^m \le c\left(\sum\sum\left(E\left(d_{rs}(\beta)\left(Z_{rs}^2 - 1\right)\right)^m\right)^{2/m}\right)^{m/2}$$

for some constant c depending on m. Hence

$$P\left\{\sum\sum d_{rs}(\beta) Z_{rs}^2 > 0\right\} \le c_1 \left[\frac{\Sigma\Sigma d_{rs}^2(\beta)}{\left(\Sigma\Sigma d_{rs}(\beta)\right)^2}\right]^{m/2}$$

for some constant c_1 depending on m.

Taking $m = 8$ and using Lemma 1 gives

$$P\left\{\sum\sum d_{rs}(\beta) Z_{rs}^2 > 0\right\} \le c_2/\beta^2,$$

where c_2 is a constant. The required result now follows since

$$\sum_{\beta=\beta_0}^{\infty} \frac{1}{\beta^2} \to 0 \quad \text{as } \beta_0 \to \infty. \qquad \square$$

Proof of Lemma 1.

Proof of (a).

$$\begin{aligned}\sum\sum c_{rs}(\beta) &= \sum\sum b_{rs}(\beta) - \beta \sum\sum b_{rs}^2(\beta) \\ &\quad - \frac{1}{n^2}\left\{\sum\sum a_{rs} b_{rs}(\beta)\right\}\left\{\sum\sum b_{rs}(\beta)\right\} \\ &= -\beta \sum\sum b_{rs}^2(\beta) + \frac{\beta}{n^2}\left(\sum\sum b_{rs}(\beta)\right)^2.\end{aligned}$$

Now

$$\sum\sum b_{rs}(\beta) \le \frac{N^2}{\pi^2}\int_0^\pi\int_0^\pi \frac{dy\,dx}{4N^4(1-\cos x)(1-\cos y)+\beta}$$

$$= \frac{N^2}{\pi\sqrt{\beta}}\int_0^\pi \frac{dx}{\left[8N^4(1-\cos x)+1\right]^{1/2}} \quad \text{[by GR (page 383)]}$$

$$= \frac{2N^2}{\pi\sqrt{\beta}\,(\beta+16N^4)^{1/2}}F\left(\frac{\pi}{2},r\right) \quad \text{[by GR (page 154)]},$$

where $r = (16N^4/(\beta + 16N^4))^{1/2}$, and F is the elliptic integral of the first kind. Also

$$\sum\sum b_{rs}^2(\beta) \ge \frac{N^2}{\pi^2}\left\{\int_0^\pi\int_0^\pi \frac{dy\,dx}{\left[4N^4(1-\cos x)(1-\cos y)+\beta\right]^2} - 2\int_0^{\pi/N}\int_0^\pi \frac{dy\,dx}{\left[4N^4(1-\cos x)(1-\cos y)+\beta\right]^2}\right\}$$

$$= \frac{N^2}{\pi\beta^{3/2}}\left\{\int_0^\pi \frac{4N^4(1-\cos x)+\beta}{\left[8N^4(1-\cos x)+\beta\right]^{3/2}}\,dx - 2\int_0^{\pi/N} \frac{4N^4(1-\cos x)+\beta}{\left[8N^4(1-\cos x)+\beta\right]^{3/2}}\,dx\right\}$$

[by GR (page 383)]

$$= \frac{N^2}{\pi\beta^{3/2}(\beta+16N^4)^{1/2}}\left\{u(\pi) - 2u\left(\frac{\pi}{N}\right)\right\},$$

where $u(x) = F[\delta(x), r] + E[\delta(x), r]$, E is the elliptic integral of the second kind, and

$$\delta(x) = \sin^{-1}\left\{\left(\frac{(16N^4+\beta)(1-\cos x)}{2\{8N^4(1-\cos x)+\beta)}\right)^{1/2}\right\}.$$

This follows using GR (page 154) and GR (page 156).

Since $u(\pi) > 0$ and $\lim_{x\to 0} u(x) = 0$, we have that for N large enough

$$\sum\sum b_{rs}^2(\beta) \ge \frac{N^2 u(\pi)}{2\pi\beta^{3/2}(\beta+16N^4)^{1/2}}.$$

Hence

$$\sum\sum c_{rs}(\beta) \le \frac{-N^2(F(\pi/2,r)+E(\pi/2,r))}{2\pi\beta^{1/2}(\beta+16N^4)^{1/2}} + \frac{\beta}{N^4}\frac{4N^4F(\pi/2,r)^2}{\pi^2\beta(\beta+16N^4)}$$

$$= \frac{-1}{\sqrt{\beta}}\phi(A),$$

where $A = \beta/N^4$ and

$$\phi(A) = \frac{F(\pi/2,r)+E(\pi/2,r)}{2\pi(A+16)^{1/2}} - \frac{4\sqrt{A}\,F(\pi/2,r)^2}{\pi^2(A+16)}$$

with

$$r = \left(\frac{16}{A+16}\right)^{1/2}.$$

It can be easily shown that $\phi(A)$ is bounded above zero for $0 \le A \le 6$. For $0 \le \delta \le 1$ we have

$$c_{rs}(\beta) - c_{rs}(\beta+\delta) = \delta c'_{rs}(\beta+\delta^*) \quad \text{for some } \delta^* \in [0,\delta]$$

and where

$$c'_{rs}(\gamma) = -2a_{rs}b^3_{rs}(\gamma) + \frac{a_{rs}}{n^2}\sum_j\sum_k(2\gamma + a_{rs} + a_{jk})b^2_{rs}(\gamma)b^2_{jk}(\gamma).$$

Hence

$$(*)\qquad \begin{aligned}|c_{rs}(\beta) - c_{rs}(\beta+\delta)| &\le 2b^2_{rs}(\beta) + \frac{1}{n^2}\sum\sum b^2_{jk}(\beta)\\ &\quad + b_{rs}(\beta)\left[\frac{1}{n^2}\sum\sum b_{jk}(\beta)\right].\end{aligned}$$

Since $d_{rs}(\beta) = c_{rs}(\beta+\delta_{rs})$ for some $\delta_{rs} \in [0,1]$ we have that

$$\left|\sum\sum d_{rs}(\beta) - \sum\sum c_{rs}(\beta)\right| \le 3\sum\sum b^2_{rs}(\beta) + \frac{[\Sigma\Sigma b_{rs}(\beta)]^2}{n^2}$$

$$= O\left[\frac{1}{\beta^{3/2}}\right].$$

Hence for N_0 large enough (a) follows. □

PROOF OF (b).

$$\sum\sum d^2_{rs}(\beta) \le 2\sum\sum c^2_{rs}(\beta) + 2\sum\sum(d_{rs}(\beta) - c_{rs}(\beta))^2$$

By $(*)$ $|d_{rs}(\beta) - c_{rs}(\beta)| \le 4$ and so

$$\sum\sum (d_{rs}(\beta) - c_{rs}(\beta))^2 \le 4 \sum\sum |d_{rs}(\beta) - c_{rs}(\beta)| = O[1/\beta^{3/2}]. \qquad \square$$

PROOF OF THEOREM 3. Since $M_0(0) = 0$,

$$\begin{aligned} P\{\tilde{\beta} \ge \beta_0\} &\le P\Big\{ \sup_{\beta \ge \beta_0} M_0(\beta) > 0 \Big\} \\ &\le \sum_{\beta=\beta_0}^{\infty} P\Big\{ \sup_{\beta \le h \le \beta+1} M_0(h) > 0 \Big\} \\ &\le \sum_{\beta=\beta_0}^{\infty} P\Big\{ \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} (\beta + 1) w_{rs}(\beta + 1) Z_{rs}^2 \\ &\qquad > - \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \log[u_{rs} w_{rs}(\beta)] \Big\}, \end{aligned}$$

where

$$u_{rs} = \pi^4 r^2 s^2,$$

$$w_{rs}(\beta) = (u_{rs} + \beta)^{-1}.$$

Using Chebyshev and Whittle as in the proof of Theorem 2 gives that for any even integer m

$$P\{\tilde{\beta} \ge \beta_0\} \le c \sum_{\beta=\beta_0}^{\infty} \left\{ \frac{\beta^2 \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} w_{rs}^2(\beta)}{\left[-\sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \log[u_{rs} w_{rs}(\beta)] - \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \beta w_{rs}(\beta) \right]^2} \right\}^{m/2},$$

where c is a constant depending on m.

Using the fact that $\log((x + 1)/x) \ge 2/(2x + 1)$ we have that

$$- \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \log[u_{rs} w_{rs}(\beta)] - \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \beta w_{rs}(\beta) \ge \frac{\beta^2}{2} \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} w_{rs}^2(\beta).$$

Hence

$$P\{\tilde{\beta} \ge \beta_0\} \le c \sum_{\beta=\beta_0}^{\infty} \left\{ \frac{1}{\beta^2 \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} w_{rs}^2(\beta)} \right\}^{m/2}.$$

Now

$$\sum_{r=1}^{\infty}\sum_{s=1}^{\infty} w_{rs}^2(\beta) \geq \int_1^{\infty}\int_1^{\infty} \frac{dy\,dx}{[\beta + \pi^4 x^2 y^2]^2}$$
$$= O\left[\frac{\log(\beta)}{\beta^{3/2}}\right].$$

Taking $m = 8$ gives the theorem. □

PROOF OF THEOREM 4. Define

$$u_{rs} = \pi^4 r^2 s^2,$$
$$w_{rs}(\beta) = (\beta + u_{rs})^{-1}$$

and

$$H(\beta) = -n^2 \log\left\{\frac{\Sigma\Sigma u_{rs} w_{rs}(\beta) Z_{rs}^2}{\Sigma\Sigma Z_{rs}^2}\right\} + \sum\sum \log[u_{rs} w_{rs}(\beta)].$$

We show separately that

(a) $$\sup_{0 \leq \beta \leq \beta_0} |M(\beta) - H(\beta)| \to_P 0 \quad \text{as } n \to \infty,$$

(b) $$\sup_{0 \leq \beta \leq \beta_0} |H(\beta) - M_0(\beta)| \to_P 0 \quad \text{as } n \to \infty.$$

PROOF OF (a). Introduce positive constants R_1 and R_2 such that

$$|\pi^2 r^2 - N^2 \lambda_r| \leq R_1 \pi^4 r^4 / N^2$$

and

$$\pi^2 r^2 \leq R_2 N^2 \lambda_r.$$

Now

$$M(\beta) - H(\beta) = B_1 + B_2,$$

where

$$B_1 = n^2\left\{\log\left[\sum\sum u_{rs} w_{rs}(\beta) Z_{rs}^2\right] - \log\left[\sum\sum a_{rs} b_{rs}(\beta) Z_{rs}^2\right]\right\}$$

and

$$B_2 = \sum\sum \left[\log[a_{rs} b_{rs}(\beta)] - \log[u_{rs} w_{rs}(\beta)]\right].$$

Using the fact that

$$|\log(x_1) - \log(x_2)| \leq \frac{|x_1 - x_2|}{\min(x_1, x_2)}$$

and the above inequalities, it is straightforward to show that $B_1 \to_P 0$ and $B_2 \to_P 0$. □

PROOF OF (b). We have that

$$M_0(\beta) - H(\beta) = c_1 + c_2,$$

where

$$c_1 = \sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \beta w_{rs}(\beta) Z_{rs}^2 + n^2 \log\left\{\frac{\sum\sum u_{rs} w_{rs}(\beta) Z_{rs}^2}{\sum\sum Z_{rs}^2}\right\}$$

and

$$c_2 = \sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \log(u_{rs} w_{rs}(\beta)) - \sum\sum \log(u_{rs} w_{rs}(\beta)).$$

Now

$$\begin{aligned} n^2 \log\left\{\frac{\sum\sum u_{rs} w_{rs}(\beta) Z_{rs}^2}{\sum\sum Z_{rs}^2}\right\} &= n^2 \log\left\{1 - \frac{\beta\sum\sum w_{rs}(\beta) Z_{rs}^2}{\sum\sum Z_{rs}^2}\right\} \\ &= n^2\beta\frac{\sum\sum w_{rs}(\beta) Z_{rs}^2}{\sum\sum Z_{rs}^2} + O_p\left(\frac{1}{n^2}\right) \\ &= \beta\sum\sum w_{rs}(\beta) Z_{rs}^2 + O_p\left(\frac{1}{n}\right) \\ &= \beta\sum_{r=1}^{\infty}\sum_{s=1}^{\infty} w_{rs}(\beta) Z_{rs}^2 + O_p\left(\frac{1}{n}\right) \end{aligned}$$

and so $c_1 \to_P 0$.

Since

$$\begin{aligned} -\sum_{r=1}^{\infty}\sum_{s=n}^{\infty} \log(u_{rs} w_{rs}(\beta)) &= \sum_{r=1}^{\infty}\sum_{s=n}^{\infty} \log\left\{1 + \frac{\beta}{\pi^4 r^2 s^2}\right\} \\ &\le \frac{\beta}{\pi^4}\left(\sum_{r=1}^{\infty}\frac{1}{r^2}\right)\left(\sum_{s=n}^{\infty}\frac{1}{s^2}\right), \end{aligned}$$

we have that $c_2 \to 0$ as $n \to \infty$ □

PROOF OF THEOREM 1. If $M_0(\tilde{\beta}) \ge M(\hat{\beta})$, then

$$\begin{aligned} M_0(\tilde{\beta}) - M(\hat{\beta}) &= M_0(\tilde{\beta}) - M(\tilde{\beta}) + M(\tilde{\beta}) - M(\hat{\beta}) \\ &\le M_0(\tilde{\beta}) - M(\tilde{\beta}). \end{aligned}$$

Hence $|M_0(\tilde{\beta}) - M(\hat{\beta})| \le |M_0(\tilde{\beta}) - M(\tilde{\beta})|$.

Similarly if $M_0(\tilde{\beta}) \le M(\hat{\beta})$, then

$$\left|M_0(\tilde{\beta}) - M(\hat{\beta})\right| \le \left|M_0(\hat{\beta}) - M(\hat{\beta})\right|.$$

Hence

$$\left|M(\hat{\beta}) - M_0(\tilde{\beta})\right| \le \left|M(\hat{\beta}) - M_0(\hat{\beta})\right| + \left|M(\tilde{\beta}) - M_0(\tilde{\beta})\right|.$$

Given $\varepsilon > 0$ we have that for each $\beta_0 > 0$ that

$$P\left\{\left|M(\hat{\beta}) - M_0(\tilde{\beta})\right| > \varepsilon\right\} \le P\{\hat{\beta} > \beta_0\} + P\{\tilde{\beta} > \beta_0\} + P\left\{\sup_{0 \le \beta \le \beta_0} |M(\beta) - M_0(\beta)| > \varepsilon/2\right\}.$$

Taking limits first as $n \to \infty$ and then as $\beta_0 \to \infty$ gives the theorem by Theorems 2, 3 and 4. □

6. Conclusion. We have proposed a new test, the W test, for testing the hypothesis that a bivariate regression function is additive. The asymptotic null distribution of W has been derived for data gathered on an equally spaced grid and we have demonstrated by simulation that the asymptotic distribution is attained for quite small sample sizes. The W test has been shown to perform well in a simulation study comparing its power with that of the two principal existing tests for additivity.

REFERENCES

BARRY, D. (1983). Nonparametric Bayesian regression. Ph.D. dissertation, Yale Univ.

BARRY, D. (1986). Nonparametric Bayesian regression. *Ann. Statist.* **14** 934–953.

BARRY, D. (1988). Asymptotic IMSE for a nonparametric Bayesian regression estimate. *Comm. Statist. A—Theory Methods* **17** 3277–3293.

BARRY, D. and HARTIGAN, J. A. (1990). An omnibus test for constant mean. *Ann. Statist.* **18** 1340–1357.

BELLMAN, R. (1970). *Introduction to Matrix Analysis.* McGraw-Hill, New York.

BUJA, A., HASTIE, T. and TIBSHIRANI, R. (1989). Linear smoothers and additive models. *Ann. Statist.* **17** 453–555.

CHEN, Z. (1987). A stepwise approach for purely periodic interaction spline models. *Comm. Statist. A—Theory Methods* **16** 877–895.

CHEN, Z. (1991). Interaction spline models and their convergence rates. *Ann. Statist.* **19** 1855–1868.

COX, D. and KOH, E. (1989). A smoothing spline based test of model adequacy in polynomial regression. *Ann. Inst. Statist. Math.* **41** 383–400.

COX, D., KOH, E., WAHBA, G. and YANDELL, B. (1988). Testing the (parametric) null model hypothesis in (semiparametric) partial and generalised spline models. *Ann. Statist.* **16** 113–119.

CRAVEN, P. and WAHBA, G. (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalised cross validation. *Numer. Math.* **31** 377–403.

FRIEDMAN, J. H. and STUETZLE, W. (1981). Projection pursuit regression. *J. Amer. Statist. Assoc.* **76** 817–823.

GRADSTEYN, I. S. and RYZHIK, I. M. (1980). *Tables of Integrals, Series and Products.* Academic, New York.

GU, C., BATES, D. M., CHEN, Z. and WAHBA, G. (1988). The computation of GCV functions through Householder tridiagonalisation with application to the fitting of interaction spline models. Technical Report 823, Dept. Statistics, Univ. Wisconsin–Madison.

HEGEMANN, V. and JOHNSON, D. E. (1976). The power of two tests for non-additivity. *J. Amer. Statist. Assoc.* **71** 945–948.

JOHNSON, D. E. and GRAYBILL, F. A. (1972). An analysis of a two-way model with interaction and no replication. *J. Amer. Statist. Assoc.* **67** 862–868.

PRAKASA RAO, B. L. S. (1983). *Nonparametric Functional Estimation.* Academic, New York.

TUKEY, J. W. (1949). One degree of freedom for non-additivity. *Biometrics* **5** 232–242.

WAHBA, G. (1978). Improper priors, spline smoothing and the problem of guarding against model errors in regression. *J. Roy. Statist. Soc. Ser. B* **40** 364–372.

WAHBA, G. (1986). Partial and interaction splines for the semiparametric estimation of functions of several variables. In *Computer Science and Statistics: Proceedings of the 18th Symposium on the Interface* 75–80. Amer. Statist. Assoc., Washington, D.C.

WHITTLE, P. (1960). Bounds for the moments of linear and quadratic forms in independent variables. *Theory Probab. Appl.* **5** 302–305.

YANAGIMOTO, T. and YANAGIMOTO, M. (1987). The use of marginal likelihood for a diagnostic test for the goodness of fit of the simple linear regression model. *Technometrics* **29** 95–107.

DEPARTMENT OF STATISTICS
UNIVERSITY COLLEGE
CORK
IRELAND