Uniform Error Bounds Involving Logspline Models

By

Charles J. Stone

Technical Report No. 130
November 1987
(revised September 1988)

Research supported in part by NSF Grant DMS-8600409.

Department of Statistics
University of California
Berkeley, California

UNIFORM ERROR BOUNDS INVOLVING LOGSPLINE MODELS [1]

Charles J. Stone

Department of Statistics
University of California, Berkeley

1. INTRODUCTION.

Splines are of increasing importance in statistical theory and methodology. In particular, Stone and Koo (1986) and Stone (1988) considered exponential families of densities in which the logarithm of the density is a spline. Such exponential families are the subject of the present paper, as are corresponding exponential response models. In each context we use an extension of a key result of de Boor (1976) to obtain a bound on the L_∞ norm of the approximation error associated with maximizing the associated expected log-likelihood.

Let Y be a real-valued random variable ranging over a compact interval $\mathcal{I}$; without loss of generality, let $\mathcal{I} = [0,1]$. Suppose that Y has a density f that is continuous and positive on $\mathcal{I}$.

Let S be a standard vector space of spline functions of a given order $q \geq 1$ on $\mathcal{I}$ (piecewise polynomials of degree $q-1$ or less that are right-continuous on $\mathcal{I}$ and continuous at 1) having finite dimension $K \geq 2$. Let $B_1, \ldots, B_K$ be a B-spline basis of S (see de Boor, 1978). Then $B_1, \ldots, B_K$ are nonnegative and sum to 1 on $\mathcal{I}$.

Let $\theta_1, \ldots, \theta_K$ be real constants. Set

$$c(\theta_1, \ldots, \theta_K) = \log\left(\int \exp\left(\sum_k \theta_k B_k(y)\right) dy\right)$$

[1] This research was supported in part by National Science Foundation Grant DMS-8600409.

and

$$f(y;\theta_1,\ldots,\theta_K) = exp\left(\sum_k \theta_k B_k(y) - c(\theta_1,\ldots,\theta_K)\right), \quad y \in \mathcal{I}.$$

This defines an exponential family of densities on $\mathcal{I}$. Observe that, for $a \in \mathbf{R}$,

$$c(\theta_1 + a,\ldots,\theta_K + a) = c(\theta_1,\ldots,\theta_K) + a$$

and hence

$$f(y;\theta_1 + a,\ldots,\theta_K + a) = f(y;\theta_1,\ldots,\theta_K), \quad y \in \mathbf{R}.$$

Consequently the exponential family fails to be identifiable. In order to make it identifiable, we require that $\theta_K = 0$.

Let Θ denote the collection of ordered $(K-1)$-tuples $\theta_1,\ldots,\theta_{K-1}$ of real numbers. For $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_{K-1}) \in \Theta$, set

$$s(y;\boldsymbol{\theta}) = \theta_1 B_1(y) + \cdots + \theta_{K-1}B_{K-1}(y), \quad y \in \mathcal{I},$$

$$C(\boldsymbol{\theta}) = \log\left(\int \exp(s(y;\boldsymbol{\theta}))dy\right),$$

and

$$f(y;\boldsymbol{\theta}) = \exp(s(y;\boldsymbol{\theta}) - C(\boldsymbol{\theta})), \quad y \in \mathcal{I}.$$

This defines an identifiable exponential family; it is referred to as a *logspline model* since $\log(f(\cdot;\boldsymbol{\theta})) \in \mathcal{S}$.

Let $Y_1 \ldots Y_n$ be independent random variables having common density f, which is not necessarily a member of the indicated logspline model. The corresponding log-likelihood function $l(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$, is defined by

$$l(\boldsymbol{\theta}) = \sum_i \log(f(Y_i;\boldsymbol{\theta})) = \sum_i [s(Y_i;\theta) - C(\boldsymbol{\theta})], \quad \boldsymbol{\theta} \in \Theta.$$

Suppose that (for given values of $Y_1,\ldots,Y_n$) the log-likelihood function has a maximizing value $\hat{\theta} \in \Theta$. Then this maximizing value is unique and is called the maximum-likelihood estimate of $\boldsymbol{\theta}$; the corresponding density $\hat{f}$ defined by $\hat{f}(y) = f(y;\hat{\theta})$ for $y \in \mathcal{I}$, is referred to as the *logspline density estimate* corresponding to the given logspline model.

The expected log-likelihood function $\lambda(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$, is defined by

$$\lambda(\boldsymbol{\theta}) = El(\boldsymbol{\theta}) = n\left[\int s(y;\boldsymbol{\theta})f(y)dy - C(\boldsymbol{\theta})\right], \quad \boldsymbol{\theta} \in \Theta.$$

It follows by a convexity argument that the expected log-likelihood function has a unique maximizing value $\boldsymbol{\theta}^* \in \Theta$. (Recall that f is a positive density on $\mathcal{I}$ and that $s(\cdot;\theta)$ is a nonconstant function for $\theta \neq 0$.) Consider the corresponding density $Q_S f$ on $\mathcal{I}$ defined by $Q_S f(y) = f(y;\theta^*)$, $y \in \mathcal{I}$. The density f belongs to the logspline model if and only if $f = Q_S f$ on $\mathcal{I}$. When f does not belong to this model, the function $f - Q_S f$ plays an important role in the analysis of the asymptotic behavior of the logspline density estimate (see Stone, 1988); roughly speaking, it acts as a bias term.

Given a real-valued function g on $\mathcal{I}$, set $\| g \|_\infty = \sup_{\mathcal{I}} | g(y) |$. Let $\mathcal{F}$ denote a family of positive densities on $\mathcal{I}$ such that the family $\{\log(f) : f \in \mathcal{F}\}$ is an equicontinuous family. Set

$$\delta_S(f) = \inf_{s \in S} \| \log(f) - s \|_\infty, \quad f \in \mathcal{F}.$$

(For an upper bound to $\delta_S(f)$ in terms of the smoothness of $\log(f)$, see Theorem XII.1 of de Boor, (1978.) In Section 4 we will obtain an inequality of the form

$$\| \log(f) - \log(Q_S f) \|_\infty \leq M \delta_S(f), \quad f \in \mathcal{F}, \tag{1}$$

where the positive constant M depends only on $\mathcal{F}$, the order of $\mathcal{S}$, and a bound on a suitable "global mesh ratio" of $\mathcal{S}$. The main point of this result is that M does not depend on $K = \dim(\mathcal{S})$. It follows from (1) that

$$\| f - Q_S f \|_\infty \leq [\exp(M \delta_S(f) - 1)] \| f \|_\infty, \quad f \in \mathcal{F}.$$

Suppose now that the distribution of Y depends on a real variable x that ranges over a compact interval $\mathcal{I}$; without loss of generality, let $\mathcal{I} = [0,1]$. Let $f(\cdot \mid x)$ denote the dependence of density of Y on x. It is supposed that $f(y \mid x)$, $x, y \in \mathcal{I}$, is a continuous and positive function.

Let $\mathcal{H}$ be a standard finite-dimensional vector space of spline functions of a given order on $\mathcal{I}$ having dimension $J \geq 1$, and let $H_1, \ldots, H_J$ be a B-spline basis of $\mathcal{H}$.

Let $\mathcal{B}$ denote the collection of $J \times (K-1)$ matrices $\beta = (\beta_{jk})$ of real numbers β_{jk}, $1 \leq j \leq J$ and $1 \leq k \leq K-1$. Let $\beta \in \mathcal{B}$. For $1 \leq k \leq K-1$, let $h_k(\cdot;\beta)$ be the real-valued function on $\mathcal{I}$ defined by

$$h_k(x;\beta) = \sum_j \beta_{jk} H_j(x), \quad x \in \mathcal{I}.$$

Set

$$\mathbf{h}(x;\beta) = (h_1(x;\beta), \ldots, h_{K-1}(x;\beta)), \quad x \in \mathcal{I}.$$

Then $\mathbf{h}(\cdot;\beta)$ is an $\mathbf{R}^{K-1}$-valued function on $\mathcal{I}$.

The logspline response model corresponding to $\mathcal{H}$ and $\mathcal{S}$ is defined by

$$f(y \mid x;\beta) = f(y;\mathbf{h}(x;\beta)) = \exp(s(y;\mathbf{h}(x;\beta)) - C(\mathbf{h}(x;\beta)))$$

for $\beta \in \mathcal{B}$ and $x, y \in \mathcal{I}$. Observe that, for $\beta \in \mathcal{B}$ and $x \in \mathcal{I}$, $f(\cdot \mid x; \beta)$ is a positive density on $\mathcal{I}$.

Let $x_1, \ldots, x_n \in \mathcal{I}$ and let $Y_1, \ldots, Y_n$ be independent random variables such that Y_i has density $f(\cdot \mid x_i)$. The corresponding log-likelihood function $l(\beta)$, $\beta \in \mathcal{B}$, is defined by

$$l(\beta) = \sum_i \log(f(Y_i \mid x_i; \beta)) = \sum_i (s(Y_i; \mathbf{h}(x_i; \beta)) - C(\mathbf{h}(x_i; \beta))), \quad \beta \in \mathcal{B}.$$

The expected log-likelihood function $\lambda(\beta)$, $\beta \in \mathcal{B}$, is defined by

$$\lambda(\beta) = El(\beta) = \sum_i \left[\int s(y; \mathbf{h}(x_i; \beta)) f(y \mid x_i) dy - C(\mathbf{h}(x_i; \beta)) \right], \quad \beta \in \mathcal{B}.$$

Suppose that $\mathcal{H}$ is identifiable from $x_1, \ldots, x_n$; that is, that if $h \in \mathcal{H}$ and $h(x_1) = \cdots = h(x_n) = 0$, then $h = 0$ on $\mathcal{I}$. Then, by a convexity argument, the expected log-likelihood function has a unique maximum $\beta^* \in \mathcal{B}$. Consider the corresponding function $Q_{\mathcal{S}} f$ on $\mathcal{I} \times \mathcal{I}$ defined by

$$Q_{\mathcal{S}} f(y \mid x) = f(y \mid x; \beta^*), \quad x, y \in \mathcal{I}.$$

Let $\mathcal{T}$ denote the tensor product of $\mathcal{H}$ and $\mathcal{S}$; that is, the vector space of real-valued functions on $\mathcal{I} \times \mathcal{I}$ spanned by functions of the form $h(x)s(y)$, $x, y \in \mathcal{I}$, as h and s range over $\mathcal{H}$ and $\mathcal{S}$ respectively. Then $\mathcal{T}$ has dimension JK, and the functions $H_j(x)B_k(y)$, $x, y \in \mathcal{I}$, $1 \le j \le J$ and $1 \le k \le K$ form a basis of $\mathcal{T}$.

Given a real-valued function g on $\mathcal{I} \times \mathcal{I}$, set $\| g \|_\infty = \sup_{\mathcal{I} \times \mathcal{I}} g(x, y)$. Let $\mathcal{F}$ denote a family of continuous and positive functions f on $\mathcal{I} \times \mathcal{I}$ such that $f(\cdot \mid x)$ is a density on $\mathcal{I}$ for $x \in \mathcal{I}$ and $\{\log(f) : f \in \mathcal{F}\}$ is an equicontinuous family of functions on $\mathcal{I} \times \mathcal{I}$. Set

$$\delta_{\mathcal{T}}(f) = \inf_{t \in \mathcal{T}} \| \log(f) - t \|_\infty, \quad f \in \mathcal{F}.$$

(For an upper bound to $\delta_{\mathcal{T}}(f)$ in terms of the smoothness of $\log(f)$, see Theorem 12.8 of Schumaker, 1981.) In Section 5 we will obtain an inequality of the form

$$\| \log(f) - \log(Q_{\mathcal{T}} f) \|_\infty \le M \delta_{\mathcal{T}}(f), \quad f \in \mathcal{F}, \tag{2}$$

where the positive constant M depends on $\mathcal{F}$, the orders of $\mathcal{H}$ and $\mathcal{S}$, bounds on the global mesh ratios of $\mathcal{H}$ and $\mathcal{S}$, and a measure of regularity of $x_1, \ldots, x_n$ that depends on $\mathcal{H}$. The main point of this result is that M does not depend on $J = \dim(\mathcal{H})$ or $K = \dim(\mathcal{S})$.

2. PRELIMINARY INEQUALITIES

The bound on the global mesh ratio for $\mathcal{S}$ described in de Boor (1976) is equivalent to a bound of the form

$$M^{-1}K^{-1} \le \int B_k(y)dy \le M_1 K^{-1}, \quad 1 \le k \le K, \tag{3}$$

where $M_1 > 1$ is a constant. Since the support of B_k is an interval having length $q \int B_k(y)dy$, where q is the order of $\mathcal{S}$, (3) can be written as a two-sided bound on this length. Under (3) there is a constant $M_2 > 1$ (depending on the order of $\mathcal{S}$) such that, for $\theta_1, \ldots, \theta_K \in \mathbf{R}$,

$$M_1^{-1} M_2^{-1} K^{-1} \sum_k \theta_k^2 \le \int \left(\sum_k \theta_k B_k(y) \right)^2 dy \le M_1 K^{-1} \sum_k \theta_k^2 \tag{4}$$

(see (7) of de Boor, 1976).

Similarly, we assume that

$$M_1^{-1} J^{-1} \le \int H_j(x)dx \le M_1 J^{-1}, \quad 1 \le j \le J. \tag{5}$$

Under (5) it can be assumed that, for $\beta_1, \ldots, \beta_J \in \mathbf{R}$,

$$M_1^{-1} M_2^{-1} J^{-1} \sum_j \beta_j^2 \le \int \left(\sum_j \beta_j H_j(x) \right)^2 dx \le M_1 J^{-1} \sum_j \beta_j^2. \tag{6}$$

For a given order q of $\mathcal{H}$, the functions in $\mathcal{H}$ are piecewise polynomials of degree $q - 1$ or less. In light of (5), a natural regularity assumption on $x_1, \ldots, x_n$ is that

$$M_3^{-1} n \int h^2(x)dx \le \sum_i h^2(x_i) \le M_3 n \int h^2(x)dx, \quad h \in \mathcal{H}, \tag{7}$$

for some constant $M_3 > 1$. It follows from (7) that $\mathcal{H}$ is identifiable from $x_1, \ldots, x_n$. It also follows from (7), by choosing M_3 larger if necessary depending on the order of $\mathcal{H}$, that

$$\sum_i H_j(x_i) \le M_3 J^{-1} n, \quad 1 \le j \le J. \tag{8}$$

(Let h denote the sum of the H_k's whose support overlaps with that of H_j; note that $H_j \le 1 = h = h^2$ on the support of H_j.)

Let ρ be a positive (Borel) function on $\mathcal{I}$ such that, for some constant $M_4 > 1$,

$$M_4^{-1} \leq \rho(y) \leq M_4, \quad y \in \mathcal{I}. \tag{9}$$

For the real-valued function g on $\mathcal{I} \times \mathcal{I}$, let $\| g \|_2$ be the nonnegative square root of

$$\| g \|^2 = \sum_i \int g^2(x_i, y)\rho(y)dy.$$

For $1 \leq j \leq J$ and $1 \leq k \leq K$, define B_{jk} on $\mathcal{I} \times \mathcal{I}$ by

$$B_{jk}(x, y) = H_j(x)B_k(y), \quad x, y \in \mathcal{I}.$$

It follows from (4), (6), (7) and (9) that, for $\beta \in \mathcal{B}$,

$$\frac{n}{M_1^2 M_2^2 M_3 M_4 JK} \sum_j \sum_k \beta_{jk}^2 \leq \left\| \sum_j \sum_k \beta_{jk} B_{jk} \right\|_2^2 \leq \frac{M_1^2 M_3 M_4 n}{JK} \sum_j \sum_k \beta_{jk}^2. \tag{10}$$

3. THE INVERSE GRAM MATRIX

Consider the $K \times K$ matrix M whose (k, l)th entry is $\int B_k(y)B_l(y)\rho(y)dy$. It follows from (4) that M is invertible. Let α_{kl} denote the (k, l)th entry of $\mathbf{M}^{-1}$. Then

$$\| \mathbf{M}^{-1} \|_\infty \leq \max_k \sum_l | \alpha_{kl} | .$$

By a slight extension of a result in de Boor (1976), there is a constant $M_8 > 1$, depending on M_1, M_2 and M_4, such that

$$\| \mathbf{M}^{-1} \|_\infty \leq M_8 K \tag{11}$$

(see the proof of (18) below). This has the following consequence.

Lemma 1. *Set* $g = \sum_k \theta_k B_k$. *Then*

$$\max_k | \theta_k | \leq M_8 K \max_k \left| \int g(y)B_k(y)\rho(y)dy \right| .$$

For real-valued functions g_1 and g_2 on $\mathcal{I} \times \mathcal{I}$ such that the norms $\| g_1 \|_2$ and $\| g_2 \|_2$ are finite, set

$$\langle g_1, g_2 \rangle = \sum_i \int g_1(x_i, y) g_2(x_i, y) \rho(y) dy.$$

Then $\| g \|_2^2 = \langle g, g \rangle$. Consider now the $JK \times JK$ matrix $\mathbf{M}$ whose $((j,k),(l,m))$th entry is the inner product $\langle B_{jk}, B_{lm} \rangle$ of B_{jk} and B_{lm}. It follows from (10) that $\mathbf{M}$ is invertible. Let α_{jklm} denote the $((j,k),(l,m))$th entry of $\mathbf{M}^{-1}$. Then

$$\| \mathbf{M}^{-1} \|_\infty = \max_{j,k} \sum_l \sum_m | \alpha_{jklm} | . \tag{12}$$

We will now imitate the elegant proof of (11) above in de Boor's paper (see also Descloux, 1972).

Set

$$f_{jk} = \sum_l \sum_m \alpha_{jklm} B_{lm}.$$

Then $\langle f_{jk}, B_{lm} \rangle$ equals 1 if $j = l$ and $k = m$ and it equals zero otherwise. Consequently,

$$0 < \| f_{jk} \|_2^2 = \alpha_{jkjk}.$$

Set $M_5 = M_1^2 M_2^2 M_3 M_4 > 1$. Then, by (10),

$$M_5^{-1} J^{-1} K^{-1} n \alpha_{jkjk}^2 \leq M_5^{-1} J^{-1} K^{-1} n \sum_l \sum_m \alpha_{jklm}^2 \leq \| f_{jk} \|_2^2 = \alpha_{jkjk}.$$

Therefore

$$\alpha_{jkjk} \leq M_5 J K n^{-1}$$

and

$$\sum_l \sum_m \alpha_{jklm}^2 \leq M_5 J K n^{-1} \alpha_{jkjk} \leq (M_5 J K n^{-1})^2. \tag{13}$$

Set $M_6 = M_1^2 M_2 M_3 M_4 > 1$.

LEMMA 2. *There is a constant $M_7 > 1$, depending on M_6, such that*

$$| \alpha_{jklm} | \leq M_5 M_6 M_7 J K M_7^{-(|j-l|+|k-m|)} n^{-1}.$$

PROOF. Let (j,k) be given and let $v, w \in \mathbf{R}$ with $v^2 + w^2 = 1$. For $c \in \mathbf{R}$, set

$$S_c = \{(l,m) : v(l-j) + w(m-k) \geq c\}$$

and

$$g_c = \sum_{S_c}\sum \alpha_{jklm} B_{lm}.$$

Let $c > 0$. Since f_{jk} is orthogonal to B_{lm} for $(l,m) \neq (j,k)$, g_c is orthogonal to f_{jk}. There is a positive constant u, depending only on the order of $\mathcal{H}$ and $\mathcal{S}$, such that if $(l,m) \in S_c$ and $(l_1, m_1) \neq S_{c-u}$, then B_{lm} and $B_{l_1 m_1}$ have disjoint support and hence are orthogonal to each other. Consequently, g_c is orthogonal to $f_{jk} - g_{c-u}$ and hence to g_{c-u}. Therefore,

$$\| g_{c-u} \|_2^2 + \| g_c \|_2^2 = \| g_{c-u} - g_c \|_2^2$$

and hence

(14) $$\| g_{c-u} \|_2^2 \leq \| g_{c-u} - g_u \|_2^2 .$$

Now

$$g_{c-u} - g_u = \sum_{S_{c-u,c}}\sum \alpha_{jklm} B_{lm},$$

where

$$S_{c-u,c} = S_{c-u} \backslash S_c = \{(l,m) : c - u \leq v(l-j) + w(m-k) < c\}.$$

We conclude from (10) and (14) that

(15) $$\sum_{S_{c-u,c}}\sum \alpha^2_{jklm} \geq M_6^{-2} \sum_{S_{c-u}}\sum \alpha^2_{jklm}, \quad c > 0.$$

Set

$$a_\nu = \sum_{S_{c+(\nu-1)u, c+\nu u}}\sum \alpha^2_{jklm}, \quad \nu = 0, 1, 2, \dots .$$

By (15),

(16) $$| a_\nu | \geq M_6^{-2}(| a_\nu | + | a_{\nu+1} | + \cdots), \quad \nu = 0, 1, 2, \dots .$$

According to Lemma 2 of de Boor (1976), (16) implies that

(17) $$| a_\nu | \leq | a_0 | M_6 |^2 (1 - M_6^{-2})^\nu, \quad \nu = 0, 1, 2, \dots .$$

By (13) and (17),

$$| a_\nu | \leq (M_5 M_6 J K n^{-1})^2 (1 - M_6^{-2})^\nu, \quad \nu = 0, 1, 2, \dots .$$

It follows by choosing v, w, and c appropriately that if

$$\nu \leq u^{-1}[(l-j)^2 + (m-k)^2]^{1/2},$$

then

$$| \alpha_{jklm} | \leq M_5 M_6 J K (1 - M_6^{-2})^{\nu/2} n^{-1}.$$

This yields the conclusion of the lemma.

Set

$$M_8 = M_5 M_6 M_7 (M_7 + 1)^2 (M_7 - 1)^{-2} > 1.$$

It follows from (12) and Lemma 2 that

$$\| \mathbf{M}^{-1} \|_\infty \leq M_8 J K n^{-1}. \tag{18}$$

This inequality has the following implication.

LEMMA 3. *Set*

$$g = \sum_j \sum_k \beta_{jk} B_{jk}.$$

Then

$$\max_{j,k} | \beta_{jk} | \leq M_8 J K n^{-1} \max_{j,k} | \langle g, B_{jk} \rangle | .$$

4. LOGSPLINE MODELS

In this section, we obtain (1). For f a positive density on I and $0 \leq a < 1$, let f_a denote the density on $\mathcal{I}$ defined by

$$f_a(y) = \frac{f^a(y)}{\int f^a(y) dy}.$$

It can be assumed that $f_a \in \mathcal{F}$ for $f \in \mathcal{F}$ and $0 \leq a < 1$. (Extend $\mathcal{F}$ if necessary.)

Choose $s \in \mathcal{S}$ and define the real-valued function g on $\mathbf{R}$ by

$$\int \exp(ts(y) - g(t)) Q_{\mathcal{S}} f(y) dy = 1.$$

Then

$$g'(0) = \int s(y) Q_{\mathcal{S}} f(y) dy.$$

Also

$$\int [\log(Q_{\mathcal{S}} f(y)) + ts(y) - g(t)] f(y) dy$$

is maximized at $t = 0$; hence

$$g'(0) = \int s(y) f(y) dy.$$

Thus

$$\int s(y)[Q_S f(y) - f(y)]dy = 0.$$

Consequently,

$$\int B_k(y)[Q_S f(y) - f(y)]dy = 0, \quad 1 \le k \le K, \tag{19}$$

or, equivalently,

$$\int B_k(y)[Q_S f(y) - f(y)]dy = 0, \quad 1 \le k \le K-1. \tag{20}$$

Formula (20) can also be written as

$$\frac{\partial C}{\partial \theta_k}(\theta^*) = \int B_k(y) f(y) dy, \quad 1 \le k \le K-1. \tag{21}$$

Let K be a fixed positive integer and let $\mathcal{S}$ otherwise vary subject to (3). Then $B_1 \ldots B_K$ depend continuously (in the L_2 norm) on the knot sequence defining $\mathcal{S}$. Thus it follows from (21) and the properties of the Hessian matrix of $C(\cdot)$ (e.g., it is negative definite) that θ^* depends continuously on $\int B_k(y) f(y) dy$, $1 \le k \le K-1$, and the knot sequence defining f.

Let $f \in \mathcal{F}$. There is an $s \in \mathcal{S}$ such that $\| \log(f) - s \|_\infty = \delta_S(f)$. Since f is a density on $\mathcal{I}$, we conclude that

$$\left| \log\left(\int \exp(s(y)) dy \right) \right| \le \delta_S(f).$$

Consequently, there is a $\bar{\theta} \in \Theta$ such that

$$\| \log(f) - \log(f(\cdot; \bar{\theta}) \|_\infty \le 2\delta_S(f). \tag{22}$$

Note that $Q_S \bar{f} = \bar{f}$, where $\bar{f} = \bar{f}(\cdot; \bar{\theta})$. Thus it follows from (22) and the continuity properties of θ^* described above that there is a positive constant M_{1K} (depending on M_1 and $\mathcal{F}$ as well as K) such that

$$\| \log(f(\cdot; \theta^*)) - \log(f(\cdot; \bar{\theta})) \|_\infty \le M_{1K} \delta_S(f)$$

and hence

$$\| \log(f) - \log(Q_S f) \|_\infty \le (M_{1K} + 2)\delta_S(f), \quad f \in \mathcal{F}. \tag{23}$$

Choose $\bar{\theta} \in \Theta$ such that (22) holds and set $\bar{f} = f(\cdot; \bar{\theta})$. Then

$$\| \log(f) - \log(\bar{f}) \|_\infty \le 2\delta_S(f). \tag{24}$$

There are constants M_9, $M_{10} > 1$, depending on $\mathcal{F}$, such that

$$\| f - \bar{f} \|_\infty \leq M_9 \delta_S(\mathcal{F}) \tag{25}$$

and

$$M_{10}^{-1} \leq \bar{f}(y) \leq M_{10}, \quad y \in \mathcal{I}. \tag{26}$$

By (3), (19) and (25),

$$\left| \int B_k(y)[Q_S f(y) - \bar{f}(y)] dy \right| \leq M_1 M_9 K^{-1} \delta_S(f), \quad 1 \leq k \leq K. \tag{27}$$

Write

$$\log(Q_S f) - \log(\bar{f}) = \sum_k \theta_k B_k$$

and set $\epsilon = \max_k | \theta_k |$. Now $\| \log(Q_S f) - \log(\bar{f}) \|_\infty \leq \epsilon$ and hence

$$\| \log(f) - \log(Q_S f) \|_\infty \leq \epsilon + 2\delta_S(f). \tag{28}$$

It follows from (viii) on Page 155 of de Boor (1978) that there is a positive constant M_{11}, depending on the order of $\mathcal{S}$, such that

$$\epsilon \leq M_{11} \| \log(Q_S f) - \log(\bar{f}) \|_\infty . \tag{29}$$

Suppose that $\epsilon \leq 1$. Since $Q_S f = \bar{f} \exp(\sum_k \theta_k B_k)$, we conclude from (26) that

$$\left\| Q_S f - \bar{f} - \bar{f} \sum_k \theta_k B_k \right\|_\infty \leq M_{10} \epsilon^2$$

and hence from (3) and (27) that, for $1 \leq k \leq K$,

$$\left| \int B_k(y) \sum_l \theta_l B_l(y) \bar{f}(y) dy \right| \leq M_1 M_9 K^{-1} \delta_S(f) + M_1 M_{10} K^{-1} \epsilon^2. \tag{30}$$

According to (26), (30) and Lemma 1, there is a constant $M_{12} > 1$, depending on M_1, M_2 and M_{10}, such that

$$\epsilon \leq M_1 M_9 M_{12} \delta_S(f) + M_1 M_{10} M_{12} \epsilon^2.$$

Suppose now that

$$M_1 M_{10} M_{12} \epsilon \leq \frac{1}{2}. \tag{31}$$

Then $\epsilon \leq 2M_1 M_9 M_{12} \delta_S(f)$ and hence, by (28),

$$\| \log(f) - \log(Q_S f) \|_\infty \leq M_{13} \delta_S(f), \tag{32}$$

where $M_{13} = 2(M_1 M_9 M_{12} + 1)$. According to (29), a sufficient condition for (31) and hence for (32) is

$$\| \log(Q_{\mathcal{S}} f) - \log(\bar{f}) \|_\infty \leq M_{14}^{-1}, \tag{33}$$

where $M_{14} = 2M_1 M_{10} M_{11} M_{12}$.

Let

$$0 < \delta < 2^{-1} M_{13}^{-1} M_{14}^{-1}.$$

There is a positive integer K_0, depending on M_1 and the order of $\mathcal{S}$, such that

$$\delta_{\mathcal{S}}(f) \leq \delta, \quad K \geq K_0 \text{ and } f \in \mathcal{F} \tag{34}$$

(see Page 167 of de Boor, 1978). Let $K \geq K_0$. Suppose that

$$\| \log(f) - \log(Q_{\mathcal{S}} f) \|_\infty \leq 2^{-1} M_{14}^{-1}. \tag{35}$$

Then (33) follows from (24), so (32) holds.

We will now verify that (35) necessarily holds for $K \geq K_0$. Suppose not. Now

$$\| \log(f_a) - \log(Q_{\mathcal{S}} f_a) \|_\infty$$

is continuous in a for $0 \leq a < 1$ and it approaches 0 as $a \to 0$. (According to an earlier argument, θ^* is continuous in a.) Thus there is a value of $a \in (0,1)$ such that

$$\| \log(f_a) - \log(Q_{\mathcal{S}} f_a) \|_\infty = 2^{-1} M_{14}^{-1}.$$

By the previous argument, (32) and (34) hold with f replaced by f_a; hence

$$\| \log(f_a) - \log(Q_{\mathcal{S}} f_a) \|_\infty \leq M_{13} \delta_{\mathcal{S}}(f_a) \leq M_{13}\delta < 2^{-1} M_{14}^{-1},$$

which yields a contradiction.

We have now shown that

$$\| \log(f) - \log(Q_{\mathcal{S}} f) \|_\infty \leq M_{13} \delta_{\mathcal{S}}(f), \quad K \geq K_0 \text{ and } f \in \mathcal{F}. \tag{36}$$

The desired inquality (1) follows from (36) together with (23) for $1 \leq K < K_0$.

5. LOGSPLINE RESPONSE MODELS

In this section, we obtain (2). For f a positive function on $\mathcal{I} \times \mathcal{I}$ such that $f(\cdot \mid x)$ is a density on $\mathcal{I}$ for each $x \in \mathcal{I}$ and for $0 < a < 1$, let f_a be defined on $\mathcal{I} \times \mathcal{I}$ by

$$f_a(y \mid x) = \frac{f^a(y \mid x)}{\int f^a(y \mid x)dy}.$$

It can be assumed that $f_a \in \mathcal{F}$ for $f \in \mathcal{F}$. (Extend $\mathcal{F}$ if necessary.)

Let $1 \le k \le K-1$. Choose $h \in \mathcal{H}$ and let $\mathbf{h}$ be the $\mathbf{R}^{K-1}$-valued function on $\mathcal{I}$ whose kth component is h and whose other components are zero. Define the real-valued function g on $\mathbf{R}$ by

$$g(t) = \sum_i \left[\int s(y; \mathbf{h}(x_i;\beta^*) + t\mathbf{h}(x_i))f(y \mid x_i)dy - C(\mathbf{h}(x_i;\beta^*) + t\mathbf{h}(x_i))\right].$$

Then

$$0 = g'(0) = \sum_i h(x_i)\left[\int B_k(y)f(y \mid x_i)dy - \frac{\partial C}{\partial \theta_k}(\mathbf{h}(x_i;\beta^*))\right].$$

Thus, for $1 \le j \le J$ and $1 \le k \le K-1$,

$$\sum_i H_j(x_i)\frac{\partial C}{\partial \theta_k}(\mathbf{h}(x_i;\beta^*)) = \sum_i H_j(x_i)\int B_k(y)f(y \mid x_i)dy\,, \tag{37}$$

which can also be written as

$$\sum_i H_j(x_i)\int B_k(y)[f(y \mid x_i) - Q_{\mathcal{T}}f(y \mid x_i)]dy = 0$$

or, equivalently, as

$$\sum_i H_j(x_i)\int B_k(y)[f(y \mid x_i) - Q_{\mathcal{T}}f(y \mid x_i)]dy = 0. \tag{38}$$

Let $f \in \mathcal{F}$. There is a $t \in \mathcal{T}$ such that $\| \log(f) - t \|_\infty = \delta_{\mathcal{T}}(f)$. Let $x \in \mathcal{I}$. Since $f(\cdot \mid x)$ is a density on $\mathcal{I}$, we conclude that

$$\left|\log\left(\int e^{t(x,y)}dy\right)\right| \le \delta_{\mathcal{T}}(f), \quad x \in \mathcal{I}.$$

Consequently, there is a $\bar\beta \in \mathcal{B}$ such that

$$\| \log(f) - \log(f(\cdot \mid \cdot;\bar\beta) \|_\infty \le 2\delta_{\mathcal{T}}(f). \tag{39}$$

Let J and K be fixed positive integers and let $\mathcal{H}, \mathcal{S}$ and $x_1 \dots x_n$ otherwise vary subject to (3), (5) and (7). It follows from (37) that there is a positive constant M_{JK} (depending on M_1, M_3 and $\mathcal{F}$ as well as J and K) such that

$$\| \log(f(\cdot \mid \cdot; \beta^*) - \log(f(\cdot \mid \cdot; \bar{\beta}) \|_\infty \leq M_{JK}\delta_{\mathcal{T}}(f). \tag{40}$$

We conclude from (39) and (40) that

$$\| \log(f) - \log(Q_{\mathcal{T}}f) \|_\infty \leq (M_{JK} + 2)\delta_{\mathcal{T}}(f), \quad f \in \mathcal{F}. \tag{41}$$

There are positive integers J_0 and K_0 and there is a positive constant M_9, depending on $\mathcal{F}, M_1 \dots M_4$ and the orders of $\mathcal{H}$ and $\mathcal{S}$ such that

$$\| \log(f) - \log(Q_{\mathcal{T}}f) \|_\infty \leq M_9\delta_{\mathcal{T}}(f), \quad J \geq J_0, \ \ K \geq K_0 \text{ and } f \in \mathcal{F}. \tag{42}$$

The argument used to prove (42) is a refinement of that used to prove (36). To start off, choose $\bar{t} \in \mathcal{T}$ such that $\| \log(f) - \bar{t} \|_\infty = \delta_{\mathcal{T}}(f)$, set

$$\bar{c}(x) = \log\left(\int \exp(\bar{t}(x, y))dy\right), \quad x \in \mathcal{I},$$

and note that

$$| \bar{c}(x) | \leq \delta_{\mathcal{T}}(f), \quad x \in \mathcal{I}.$$

Define $\bar{f}$ on $\mathcal{I} \times \mathcal{I}$ by $\bar{f}(y \mid x) = \exp(\bar{t}(x, y) - \bar{c}(x))$. Then

$$\| \log(f) - \log(\bar{f}) \|_\infty \leq 2\delta_{\mathcal{T}}(f).$$

There are constants M_{10}, $M_{11} > 1$, depending on $\mathcal{F}$, such that

$$\| f - \bar{f} \|_\infty \leq M_{10}\delta_{\mathcal{T}}(f) \tag{43}$$

and

$$M_{11}^{-1} \leq \bar{f}(y \mid x) \leq M_{11}, \quad x, y \in \mathcal{I}.$$

By (3), (8), (38) and (43),

$$\left| \sum_i H_j(x_i) \int B_k(y)[Q_{\mathcal{T}}f(y \mid x_i) - \bar{f}(y \mid x_i)]dy \right| \leq \frac{M_1 M_3 M_{10}}{JK} n\delta_{\mathcal{T}}(f)$$

for $1 \leq j \leq J$ and $1 \leq k \leq K$.

Write

$$\log(Q_{\mathcal{T}}f(y \mid x)) = t^*(x, y) - c^*(x), \quad x, y \in \mathcal{I},$$

where $t^* \in \mathcal{T}$, and set $t = t^* - \bar{t}$. Then

$$Q_{\mathcal{T}}f(y \mid x) = \exp(t(x, y) + \bar{c}(x) - c^*(x))\bar{f}(y \mid x), \quad x, y \in \mathcal{I},$$

$$
\begin{aligned}
c^*(x) &= \log\left(\int \exp(t(x,y)+\bar{c}(x))\bar{f}(y \mid x)dy\right) \\
&= \log\left((1+\int [\exp(t(x,y)+\bar{c}(x))-1]\bar{f}(y \mid x)dy)\right)
\end{aligned}
$$

for $x \in \mathcal{I}$, and

$$
Q_{\mathcal{T}} f(y \mid x) - \bar{f}(y \mid x) = [\exp(t(x,y)+\bar{c}(x)-c^*(x))-1]\bar{f}(y \mid x), \quad x, y \in \mathcal{I}.
$$

Thus

$$
c^*(x) - \bar{c}(x) \approx \int t(x,y)\bar{f}(y \mid x)dy, \quad x \in \mathcal{I},
$$

and hence

$$
(44) \quad Q_{\mathcal{T}} f(y \mid x) - \bar{f}(y \mid x) \approx \left[t(x,y) - \int t(x,y)\bar{f}(y \mid x)dy\right]\bar{f}(y \mid x)
$$

for $x, y \in \mathcal{I}$.

Write

$$
t(x,y) = \sum_j \sum_k \beta_{jk} H_j(x) B_k(y), \quad x, y \in \mathcal{I}.
$$

It follows by a double application of (viii) on Page 155 of de Boor (1978) that there is a positive constant M_{12}, depending on the order of $\mathcal{H}$ and $\mathcal{S}$, such that

$$
\max_{j,k} \mid \beta_{jk} \mid \leq M_{12} \parallel t \parallel_\infty .
$$

Choose $\eta > 0$. Now

$$
\int t(x,y)\bar{f}(y \mid x)dy = \sum_k \int B_k(y) \sum_j \beta_{jk} H_j(x) \bar{f}(y \mid x)dy.
$$

Choose x_j in the support of H_j. Define $h \in \mathcal{H}$ by

$$
\begin{aligned}
h(x) &= \sum_k \int B_k(y) \sum_j \beta_{jk} H_j(x) \bar{f}(y \mid x_j)dy \\
&= \sum_j H_j(x) \sum_k \beta_{jk} \int B_k(y)\bar{f}(y \mid x_j)dy.
\end{aligned}
$$

There is a positive integer J_0, depending on M_1, M_{12} and $\mathcal{F}$ such that

$$
\left|\int t(x,y)\bar{f}(y \mid x)dy - h(x)\right| \leq \eta \parallel t \parallel_\infty, \quad J \geq J_0 \text{ and } x \in \mathcal{I}.
$$

After replacing $t^*(x,y)$ by $t^*(x,y)-h(x)$ and replacing $c^*(x)$ by $c^*(x)-h(x)$, we have that

$$\left|\int t(x,y)\bar{f}(y \mid x)dy\right| \leq \eta \parallel t \parallel_\infty, \quad J \geq J_0 \text{ and } x \in \mathcal{I}. \tag{45}$$

The argument used to prove (42) from (44) and (45) is similar to that used to prove (36), except that Lemma 3 is used instead of Lemma 1 and Theorem 12.8 of Schumaker (1981) is used instead of Page 167 of de Boor (1978).

Next it will be shown that, for each positive integer K, there is a positive integer J_0 and there is a positive constant M_{13}, both depending on $\mathcal{F}$, $M_1, \ldots, M_4$ and the order of $\mathcal{H}$ and $\mathcal{S}$, such that

$$\parallel \log(f) - \log(Q_{\mathcal{T}} f) \parallel_\infty \leq M_{13}\delta_{\mathcal{F}}(f), \quad J \geq J_0 \text{ and } f \in \mathcal{F}. \tag{46}$$

To this end, write

$$Q_{\mathcal{S}} f(y \mid x) = \exp\left(\sum_k \theta_k(x)B_k(y) - c(x)\right), \quad x, y \in \mathcal{I}.$$

From (21) we conclude that (as f varies over $\mathcal{F}$, etc.) the resulting functions $\theta_k(\cdot)$, $1 \leq k \leq K-1$, are uniformly bounded and equicontinuous, and there is a positive constant M_{14} such that

$$\max_{1 \leq k \leq K-1} \delta_{\mathcal{H}}(\theta_k(\cdot)) \leq M_{14}\delta_{\mathcal{T}}(f). \tag{47}$$

Observe that

$$\max_{1 \leq k \leq K-1} \delta_{\mathcal{H}}(\theta_k(\cdot))$$

can be made arbitrary small by making J sufficiently large (see Page 167 of de Boor, 1978). According to (1), there is a positive constant M_{15} such that

$$\left|\log(f(y \mid x)) - \left(\sum_k \theta_k(x)B_k(y) - c(x)\right)\right| \leq M_{15}\delta_{\mathcal{T}}(f), \quad x, y \in \mathcal{I}. \tag{48}$$

It follows from (19) that

$$\int B_k(y)\left[\exp\left(\sum_m \theta_m(x)B_m(y) - c(x)\right) - f(y \mid x)\right] dy = 0$$

for $x \in \mathcal{I}$ and $1 \leq k \leq K$ and hence that

$$\sum_i H_j(x_i) \int B_k(y)\left[\exp\left(\sum_m \theta_m(x_i)B_m(y) - c(x_i)\right) - f(y \mid x_i)\right] dy = 0$$

for $1 \leq j \leq J$ and $1 \leq k \leq K$. Thus we conclude from (38) that

$$\sum_k H_j(x_i) \int B_k(y) \left[\exp\left(\sum_m \theta_m(x_i) B_m(y) - c(x_i) \right) - Q_{\mathcal{T}} f(y \mid x_i) \right] dy = 0$$

for $1 \leq j \leq J$ and $1 \leq k \leq K$.

For $1 \leq k \leq K-1$, choose $\bar{h}_k \in \mathcal{H}$ such that

$$| \theta_k(x) - \bar{h}_k(x) | = \delta_{\mathcal{H}}(\theta_k(\cdot)), \quad x \in \mathcal{I}.$$

Set

$$\bar{c}(x) = \log\left(\int \exp\left(\sum_k \bar{h}_k(x) B_k(y) \right) dy \right), \quad x \in \mathcal{I},$$

and define $\bar{f}$ on $\mathcal{I} \times \mathcal{I}$ by

$$\bar{f}(y \mid x) = \exp\left(\sum_k \bar{h}_k(x) B_k(y) - \bar{c}(x) \right).$$

Write

$$Q_{\mathcal{T}} f(y \mid x) = \exp\left(\sum_k h^*(x) B_k(y) - c^*(x) \right), \quad x, y \in \mathcal{I},$$

where $h^* \in \mathcal{H}$ for $1 \leq k \leq K-1$. It now follows by arguing as in the proofs of (36) and (42) that there is a positive constant M_{16} such that

$$| \theta_k(x) - h^*(x) | \leq M_{16} \max_{1 \leq k \leq K-1} \delta_{\mathcal{H}}(\theta_k(\cdot)), \quad 1 \leq k \leq K-1 \text{ and } x \in \mathcal{I}.$$

Thus there is a positive constant M_{17} such that

$$\left| \log(Q_{\mathcal{T}} f(y \mid x)) - \left(\sum_k \theta_k(x) B_k(y) - c(x) \right) \right| \leq M_{17} \max_{1 \leq k \leq K-1} \delta_{\mathcal{H}}(\theta_k(\cdot)). \tag{49}$$

The desired result (46) follows from (47)-(49).

Finally it will be shown that, for each positive integer J, there is a positive integer K_0 and there is a positive constant M_{18}, both depending on $\mathcal{F}$, $M_1, \ldots, M_4$ and the order of $\mathcal{H}$ and $\mathcal{S}$, such that

$$\| \log(f) - \log(Q_{\mathcal{T}} f) \|_\infty \leq M_{18} \delta_{\mathcal{T}}(f), \quad K \geq K_0 \text{ and } f \in \mathcal{F}. \tag{50}$$

To this end, let $\beta_1(\cdot), \ldots, \beta_J(\cdot)$ be the real-valued functions on $\mathcal{I}$ such that

$$\sum_i \left[\log(f(y \mid x_i)) - \sum_j \beta_j(y) H_j(x_i) \right]^2$$

minimizes

$$\sum_i \left[\log(f(y \mid x_i)) - \sum_j \beta_j H_j(x_i) \right]^2$$

for $y \in \mathcal{I}$. It follows from the appropriate analog of Lemma 2 that, as f varies over $\mathcal{F}$, etc., the resulting functions $\beta_1(\cdot), \cdots, \beta_J(\cdot)$ are uniformly bounded and equicontinuous, that there is a positive constant M_{19} such that

$$\max_{1 \le j \le J} \delta_S(\beta_j(\cdot)) \le M_{19} \delta_T(f), \tag{51}$$

and that there is a positive constant M_{20} such that

$$\left| \log(f(y \mid x)) - \sum_j \beta_j(y) H_j(x) \right| \le M_{20} \delta_T(f), \quad x, y \in \mathcal{I}. \tag{52}$$

Observe that

$$\max_{1 \le j \le J} \delta_S\left(\beta_j(\cdot)\right)$$

can be made arbitrarily small by making K sufficiently large. For $1 \le j \le J$ choose $\bar{s}_j \in \mathcal{S}$ such that

$$\mid \beta_j(y) - \bar{s}_j(y) \mid = \delta_S(\beta_j(\cdot)), \quad y \in \mathcal{I}. \tag{53}$$

Set

$$\bar{c}(x) = \log\left(\int \exp\left(\sum_j H_j(x) \bar{s}_j(y) dy \right) \right), \quad x \in \mathcal{I}.$$

There is a constant M_{21} such that

$$\mid \bar{c}(x) \mid \le M_{21} \delta_T(f), \quad x \in \mathcal{I}. \tag{54}$$

Define $\bar{f}$ on $\mathcal{I} \times \mathcal{I}$ by $\bar{f}(y \mid x) = \exp(\sum_j H_j(x) \bar{s}_j(y) - \bar{c}(x))$. Write

$$Q_T f(y \mid x) = \exp\left(\sum_j H_j(x) s_j^*(y) - c^*(x) \right), \quad x, y \in \mathcal{I},$$

where $s^* \in \mathcal{S}$ for $1 \le j \le J$. It follows as in the proofs of (36), (42) and (49) that there is a positive constant M_{22} such that

$$\mid \log(Q_T f(y \mid x)) - \log(\bar{f}(y \mid x)) \mid \le M_{22} \max_{1 \le j \le J} \delta_S(\beta_j(\cdot)). \tag{55}$$

The desired result (50) follows from (51)-(55).

Inequality (2) follows from (41), (42), (46), and (50).

REFERENCES

de Boor, C. (1976), A bound on the L_∞-norm of the L_2-approximation by splines in terms of a global mesh ratio, *Mathematics of Computation*, **30**, 765-771.

de Boor, C. (1978), *A Practical Guide to Splines*, Springer-Verlag, New York.

Descloux, J. (1972), On finite element matrices, *SIAM Journal on Numerical Analysis*, **9**, 260-265.

Schumaker, L. L. (1981), *Spline Functions: Basic Theory*, Wiley, New York.

Stone, C.J. (1988), Large-sample inference for logspline models, Technical Report No. 171, Department of Statistics, University of California, Berkeley.

Stone C.J. and Koo, C.-Y. (1986), Logspline density estimation, *Contemporary Mathematics*, **59**, 1-15, American Mathematical Society, Providence.

TECHNICAL REPORTS

Statistics Department

University of California, Berkeley

1. BREIMAN, L. and FREEDMAN, D. (Nov. 1981, revised Feb. 1982). How many variables should be entered in a regression equation? Jour. Amer. Statist. Assoc., March 1983, 78, No. 381, 131-136.

2. BRILLINGER, D. R. (Jan. 1982). Some contrasting examples of the time and frequency domain approaches to time series analysis. Time Series Methods in Hydrosciences, (A. H. El-Shaarawi and S. R. Esterby, eds.) Elsevier Scientific Publishing Co., Amsterdam, 1982, pp. 1-15.

3. DOKSUM, K. A. (Jan. 1982). On the performance of estimates in proportional hazard and log-linear models. Survival Analysis, (John Crowley and Richard A. Johnson, eds.) IMS Lecture Notes - Monograph Series, (Shanti S. Gupta, series ed.) 1982, 74-84.

4. BICKEL, P. J. and BREIMAN, L. (Feb. 1982). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. Ann. Prob., Feb. 1982, 11. No. 1, 185-214.

5. BRILLINGER, D. R. and TUKEY, J. W. (March 1982). Spectrum estimation and system identification relying on a Fourier transform. The Collected Works of J. W. Tukey, vol. 2, Wadsworth, 1985, 1001-1141.

6. BERAN, R. (May 1982). Jackknife approximation to bootstrap estimates. Ann. Statist., March 1984, 12 No. 1, 101-118.

7. BICKEL, P. J. and FREEDMAN, D. A. (June 1982). Bootstrapping regression models with many parameters. Lehmann Festschrift, (P. J. Bickel, K. Doksum and J. L. Hodges, Jr., eds.) Wadsworth Press, Belmont, 1983, 28-48.

8. BICKEL, P. J. and COLLINS, J. (March 1982). Minimizing Fisher information over mixtures of distributions. Sankhyā, 1983, 45, Series A, Pt. 1, 1-19.

9. BREIMAN, L. and FRIEDMAN, J. (July 1982). Estimating optimal transformations for multiple regression and correlation.

10. FREEDMAN, D. A. and PETERS, S. (July 1982, revised Aug. 1983). Bootstrapping a regression equation: some empirical results. JASA, 1984, 79, 97-106.

11. EATON, M. L. and FREEDMAN, D. A. (Sept. 1982). A remark on adjusting for covariates in multiple regression.

12. BICKEL, P. J. (April 1982). Minimax estimation of the mean of a mean of a normal distribution subject to doing well at a point. Recent Advances in Statistics, Academic Press, 1983.

14. FREEDMAN, D. A., ROTHENBERG, T. and SUTCH, R. (Oct. 1982). A review of a residential energy end use model.

15. BRILLINGER, D. and PREISLER, H. (Nov. 1982). Maximum likelihood estimation in a latent variable problem. Studies in Econometrics, Time Series, and Multivariate Statistics, (eds. S. Karlin, T. Amemiya, L. A. Goodman). Academic Press, New York, 1983, pp. 31-65.

16. BICKEL, P. J. (Nov. 1982). Robust regression based on infinitesimal neighborhoods. Ann. Statist., Dec. 1984, 12, 1349-1368.

17. DRAPER, D. C. (Feb. 1983). Rank-based robust analysis of linear models. I. Exposition and review.

18. DRAPER, D. C. (Feb 1983). Rank-based robust inference in regression models with several observations per cell.

19. FREEDMAN, D. A. and FIENBERG, S. (Feb. 1983, revised April 1983). Statistics and the scientific method, Comments on and reactions to Freedman, A rejoinder to Fienberg's comments. Springer New York 1985 Cohort Analysis in Social Research, (W. M. Mason and S. E. Fienberg, eds.).

20. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Jan. 1984). Using the bootstrap to evaluate forecasting equations. J. of Forecasting. 1985, Vol. 4, 251-262.

21. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Aug. 1983). Bootstrapping an econometric model: some empirical results. JBES, 1985, 2, 150-158.

22. FREEDMAN, D. A. (March 1983). Structural-equation models: a case study.

23. DAGGETT, R. S. and FREEDMAN, D. (April 1983, revised Sept. 1983). Econometrics and the law: a case study in the proof of antitrust damages. Proc. of the Berkeley Conference, in honor of Jerzy Neyman and Jack Kiefer. Vol I pp. 123-172. (L. Le Cam, R. Olshen eds.) Wadsworth, 1985.

24. DOKSUM, K. and YANDELL, B. (April 1983). Tests for exponentiality. Handbook of Statistics, (P. R. Krishnaiah and P. K. Sen, eds.) 4, 1984.

25. FREEDMAN, D. A. (May 1983). Comments on a paper by Markus.

26. FREEDMAN, D. (Oct. 1983, revised March 1984). On bootstrapping two-stage least-squares estimates in stationary linear models. Ann. Statist., 1984, 12, 827-842.

27. DOKSUM, K. A. (Dec. 1983). An extension of partial likelihood methods for proportional hazard models to general transformation models. Ann. Statist., 1987, 15, 325-345.

28. BICKEL, P. J., GOETZE, F. and VAN ZWET, W. R. (Jan. 1984). A simple analysis of third order efficiency of estimate Proc. of the Neyman-Kiefer Conference, (L. Le Cam, ed.) Wadsworth, 1985.

29. BICKEL, P. J. and FREEDMAN, D. A. Asymptotic normality and the bootstrap in stratified sampling. Ann. Statist 12 470-482.

30. FREEDMAN, D. A. (Jan. 1984). The mean vs. the median: a case study in 4-R Act litigation. JBES. 1985 Vol 3 pp. 1-13.

31. STONE, C. J. (Feb. 1984). An asymptotically optimal window selection rule for kernel density estimates. Ann. Statist., Dec. 1984, 12, 1285-1297.

32. BREIMAN, L. (May 1984). Nail finders, edifices, and Oz.

33. STONE, C. J. (Oct. 1984). Additive regression and other nonparametric models. Ann. Statist., 1985, 13, 689-705.

34. STONE, C. J. (June 1984). An asymptotically optimal histogram selection rule. Proc. of the Berkeley Conf. in Honor of Jerzy Neyman and Jack Kiefer (L. Le Cam and R. A. Olshen, eds.), II, 513-520.

35. FREEDMAN, D. A. and NAVIDI, W. C. (Sept. 1984, revised Jan. 1985). Regression models for adjusting the 1980 Census. Statistical Science. Feb 1986, Vol. 1, No. 1, 3-39.

36. FREEDMAN, D. A. (Sept. 1984, revised Nov. 1984). De Finetti's theorem in continuous time.

37. DIACONIS, P. and FREEDMAN, D. (Oct. 1984). An elementary proof of Stirling's formula. Amer. Math Monthly. Feb 1986, Vol. 93, No. 2, 123-125.

38. LE CAM, L. (Nov. 1984). Sur l'approximation de familles de mesures par des familles Gaussiennes. Ann. Inst. Henri Poincaré, 1985, 21, 225-287.

39. DIACONIS, P. and FREEDMAN, D. A. (Nov. 1984). A note on weak star uniformities.

40. BREIMAN, L. and IHAKA, R. (Dec. 1984). Nonlinear discriminant analysis via SCALING and ACE.

41. STONE, C. J. (Jan. 1985). The dimensionality reduction principle for generalized additive models.

42. LE CAM, L. (Jan. 1985). On the normal approximation for sums of independent variables.

43. BICKEL, P. J. and YAHAV, J. A. (1985). On estimating the number of unseen species: how many executions were there?

44. BRILLINGER, D. R. (1985). The natural variability of vital rates and associated statistics. Biometrics, to appear.

45. BRILLINGER, D. R. (1985). Fourier inference: some methods for the analysis of array and nonGaussian series data. Water Resources Bulletin, 1985, 21, 743-756.

46. BREIMAN, L. and STONE, C. J. (1985). Broad spectrum estimates and confidence intervals for tail quantiles.

47. DABROWSKA, D. M. and DOKSUM, K. A. (1985, revised March 1987). Partial likelihood in transformation models with censored data.

48. HAYCOCK, K. A. and BRILLINGER, D. R. (November 1985). LIBDRB: A subroutine library for elementary time series analysis.

49. BRILLINGER, D. R. (October 1985). Fitting cosines: some procedures and some physical examples. Joshi Festschrift, 1986. D. Reidel.

50. BRILLINGER, D. R. (November 1985). What do seismology and neurophysiology have in common? - Statistics! Comptes Rendus Math. Rep. Acad. Sci. Canada. January, 1986.

51. COX, D. D. and O'SULLIVAN, F. (October 1985). Analysis of penalized likelihood-type estimators with application to generalized smoothing in Sobolev Spaces.

52. O'SULLIVAN, F. (November 1985). A practical perspective on ill-posed inverse problems: A review with some new developments. To appear in Journal of Statistical Science.

53. LE CAM, L. and YANG, G. L. (November 1985, revised March 1987). On the preservation of local asymptotic normality under information loss.

54. BLACKWELL, D. (November 1985). Approximate normality of large products.

55. FREEDMAN, D. A. (June 1987). As others see us: A case study in path analysis. Journal of Educational Statistics. 12, 101-128.

56. LE CAM, L. and YANG, G. L. (January 1986). Replaced by No. 68.

57. LE CAM, L. (February 1986). On the Bernstein - von Mises theorem.

58. O'SULLIVAN, F. (January 1986). Estimation of Densities and Hazards by the Method of Penalized likelihood.

59. ALDOUS, D. and DIACONIS, P. (February 1986). Strong Uniform Times and Finite Random Walks.

60. ALDOUS, D. (March 1986). On the Markov Chain simulation Method for Uniform Combinatorial Distributions and Simulated Annealing.

61. CHENG, C-S. (April 1986). An Optimization Problem with Applications to Optimal Design Theory.

62. CHENG, C-S., MAJUMDAR, D., STUFKEN, J. & TURE, T. E. (May 1986, revised Jan 1987). Optimal step type design for comparing test treatments with a control.

63. CHENG, C-S. (May 1986, revised Jan. 1987). An Application of the Kiefer-Wolfowitz Equivalence Theorem.

64. O'SULLIVAN, F. (May 1986). Nonparametric Estimation in the Cox Proportional Hazards Model.

65. ALDOUS, D. (JUNE 1986). Finite-Time Implications of Relaxation Times for Stochastically Monotone Processes.

66. PITMAN, J. (JULY 1986, revised November 1986). Stationary Excursions.

67. DABROWSKA, D. and DOKSUM, K. (July 1986, revised November 1986). Estimates and confidence intervals for median and mean life in the proportional hazard model with censored data.

68. LE CAM, L. and YANG, G.L. (July 1986). Distinguished Statistics, Loss of information and a theorem of Robert B. Davies (Fourth edition).

69. STONE, C.J. (July 1986). Asymptotic properties of logspline density estimation.

71. BICKEL, P.J. and YAHAV, J.A. (July 1986). Richardson Extrapolation and the Bootstrap.

72. LEHMANN, E.L. (July 1986). Statistics - an overview.

73. STONE, C.J. (August 1986). A nonparametric framework for statistical modelling.

74. BIANE, PH. and YOR, M. (August 1986). A relation between Lévy's stochastic area formula, Legendre polynomial, and some continued fractions of Gauss.

75. LEHMANN, E.L. (August 1986, revised July 1987). Comparing Location Experiments.

76. O'SULLIVAN, F. (September 1986). Relative risk estimation.

77. O'SULLIVAN, F. (September 1986). Deconvolution of episodic hormone data.

78. PITMAN, J. & YOR, M. (September 1987). Further asymptotic laws of planar Brownian motion.

79. FREEDMAN, D.A. & ZEISEL, H. (November 1986). From mouse to man: The quantitative assessment of cancer risks. To appear in Statistical Science.

80. BRILLINGER, D.R. (October 1986). Maximum likelihood analysis of spike trains of interacting nerve cells.

81. DABROWSKA, D.M. (November 1986). Nonparametric regression with censored survival time data.

82. DOKSUM, K.J. and LO, A.Y. (November 1986). Consistent and robust Bayes Procedures for Location based on Partial Information.

83. DABROWSKA, D.M., DOKSUM, K.A. and MIURA, R. (November 1986). Rank estimates in a class of semiparametric two-sample models.

84. BRILLINGER, D. (December 1986). Some statistical methods for random process data from seismology and neurophysiology.

85. DIACONIS, P. and FREEDMAN, D. (December 1986). A dozen de Finetti-style results in search of a theory. Ann. Inst. Henri Poincaré, 1987, 23, 397-423.

86. DABROWSKA, D.M. (January 1987). Uniform consistency of nearest neighbour and kernel conditional Kaplan - Meier estimates.

87. FREEDMAN, D.A., NAVIDI, W. and PETERS, S.C. (February 1987). On the impact of variable selection in fitting regression equations.

88. ALDOUS, D. (February 1987, revised April 1987). Hashing with linear probing, under non-uniform probabilities.

89. DABROWSKA, D.M. and DOKSUM, K.A. (March 1987, revised January 1988). Estimating and testing in a two sample generalized odds rate model.

90. DABROWSKA, D.M. (March 1987). Rank tests for matched pair experiments with censored data.

91. DIACONIS, P and FREEDMAN, D.A. (April 1988). Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. To appear in the Journal of Applied Probability.

92. DABROWSKA, D.M. (April 1987, revised September 1987). Kaplan-Meier estimate on the plane.

92a. ALDOUS, D. (April 1987). The Harmonic mean formula for probabilities of Unions: Applications to sparse random graphs.

93. DABROWSKA, D.M. (June 1987, revised Feb 1988). Nonparametric quantile regression with censored data.

94. DONOHO, D.L. & STARK, P.B. (June 1987). Uncertainty principles and signal recovery.

95. CANCELLED

96. BRILLINGER, D.R. (June 1987). Some examples of the statistical analysis of seismological data. To appear in *Proceedings, Centennial Anniversary Symposium, Seismographic Stations, University of California, Berkeley.*

97. FREEDMAN, D.A. and NAVIDI, W. (June 1987). On the multi-stage model for carcinogenesis. To appear in Environmental Health Perspectives.

98. O'SULLIVAN, F. and WONG, T. (June 1987). Determining a function diffusion coefficient in the heat equation.

99. O'SULLIVAN, F. (June 1987). Constrained non-linear regularization with application to some system identification problems.

100. LE CAM, L. (July 1987, revised Nov 1987). On the standard asymptotic confidence ellipsoids of Wald.

101. DONOHO, D.L. and LIU, R.C. (July 1987). Pathologies of some minimum distance estimators. Annals of Statistics, June, 1988.

102. BRILLINGER, D.R., DOWNING, K.H. and GLAESER, R.M. (July 1987). Some statistical aspects of low-dose electron imaging of crystals.

103. LE CAM, L. (August 1987). Harald Cramér and sums of independent random variables.

104. DONOHO, A.W., DONOHO, D.L. and GASKO, M. (August 1987). Macspin: Dynamic graphics on a desktop computer. IEEE Computer Graphics and applications, June, 1988.

105. DONOHO, D.L. and LIU, R.C. (August 1987). On minimax estimation of linear functionals.

106. DABROWSKA, D.M. (August 1987). Kaplan-Meier estimate on the plane: weak convergence, LIL and the bootstrap.

107. CHENG, C-S. (August 1987). Some orthogonal main-effect plans for asymmetrical factorials.

108. CHENG, C-S. and JACROUX, M. (August 1987). On the construction of trend-free run orders of two-level factorial designs.

109. KLASS, M.J. (August 1987). Maximizing $E \max_{1 \le k \le n} S_k^+ / ES_n^+$: A prophet inequality for sums of I.I.D. mean zero variates.

110. DONOHO, D.L. and LIU, R.C. (August 1987). The "automatic" robustness of minimum distance functionals. Annals of Statistics, June, 1988.

111. BICKEL, P.J. and GHOSH, J.K. (August 1987, revised June 1988). A decomposition for the likelihood ratio statistic and the Bartlett correction — a Bayesian argument.

112. BURDZY, K., PITMAN, J.W. and YOR, M. (September 1987). Some asymptotic laws for crossings and excursions.

113. ADHIKARI, A. and PITMAN, J. (September 1987). The shortest planar arc of width 1.

114. RITOV, Y. (September 1987). Estimation in a linear regression model with censored data.

115. BICKEL, P.J. and RITOV, Y. (Sept. 1987, revised Aug 1988). Large sample theory of estimation in biased sampling regression models I.

116. RITOV, Y. and BICKEL, P.J. (Sept.1987, revised Aug. 1988). Unachievable information bounds in non and semiparametric models.

117. RITOV, Y. (October 1987). On the convergence of a maximal correlation algorithm with alternating projections.

118. ALDOUS, D.J. (October 1987). Meeting times for independent Markov chains.

119. HESSE, C.H. (October 1987). An asymptotic expansion for the mean of the passage-time distribution of integrated Brownian Motion.

120. DONOHO, D. and LIU, R. (October 1987, revised March 1988). Geometrizing rates of convergence, II.

121. BRILLINGER, D.R. (October 1987). Estimating the chances of large earthquakes by radiocarbon dating and statistical modelling. To appear in *Statistics a Guide to the Unknown.*

122. ALDOUS, D., FLANNERY, B. and PALACIOS, J.L. (November 1987). Two applications of urn processes: The fringe analysis of search trees and the simulation of quasi-stationary distributions of Markov chains.

123. DONOHO, D.L., MACGIBBON, B. and LIU, R.C. (Nov.1987, revised July 1988). Minimax risk for hyperrectangles.

124. ALDOUS, D. (November 1987). Stopping times and tightness II.

125. HESSE, C.H. (November 1987). The present state of a stochastic model for sedimentation.

126. DALANG, R.C. (December 1987, revised June 1988). Optimal stopping of two-parameter processes on nonstandard probability spaces.

127. Same as No. 133.

128. DONOHO, D. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean II.

129. SMITH, D.L. (December 1987). Exponential bounds in Vapnik-Červonenkis classes of index 1.

130. STONE, C.J. (Nov.1987, revised Sept. 1988). Uniform error bounds involving logspline models.

131. Same as No. 140

132. HESSE, C.H. (December 1987). A Bahadur - Type representation for empirical quantiles of a large class of stationary, possibly infinite - variance, linear processes

133. DONOHO, D.L. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean, I.

134. DUBINS, L.E. and SCHWARZ, G. (December 1987). A sharp inequality for martingales and stopping-times.

135. FREEDMAN, D.A. and NAVIDI, W. (December 1987). On the risk of lung cancer for ex-smokers.

136. LE CAM, L. (January 1988). On some stochastic models of the effects of radiation on cell survival.

137. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the uniform consistency of Bayes estimates for multinomial probabilities.

137a. DONOHO, D.L. and LIU, R.C. (1987). Geometrizing rates of convergence, I.

138. DONOHO, D.L. and LIU, R.C. (January 1988). Geometrizing rates of convergence, III.

139. BERAN, R. (January 1988). Refining simultaneous confidence sets.

140. HESSE, C.H. (December 1987). Numerical and statistical aspects of neural networks.

141. BRILLINGER, D.R. (January 1988). Two reports on trend analysis: a) An Elementary Trend Analysis of Rio Negro Levels at Manaus, 1903-1985 b) Consistent Detection of a Monotonic Trend Superposed on a Stationary Time Series

142. DONOHO, D.L. (Jan. 1985, revised Jan. 1988). One-sided inference about functionals of a density.

143. DALANG, R.C. (February 1988). Randomization in the two-armed bandit problem.

144. DABROWSKA, D.M., DOKSUM, K.A. and SONG, J.K. (February 1988). Graphical comparisons of cumulative hazards for two populations.

145. ALDOUS, D.J. (February 1988). Lower bounds for covering times for reversible Markov Chains and random walks on graphs.

146. BICKEL, P.J. and RITOV, Y. (Feb.1988, revised August 1988). Estimating integrated squared density derivatives.

147. STARK, P.B. (March 1988). Strict bounds and applications.

148. DONOHO, D.L. and STARK, P.B. (March 1988). Rearrangements and smoothing.

149. NOLAN, D. (March 1988). Asymptotics for a multivariate location estimator.

150. SEILLIER, F. (March 1988). Sequential probability forecasts and the probability integral transform.

151. NOLAN, D. (March 1988). Limit theorems for a random convex set.

152. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On a theorem of Kuchler and Lauritzen.

153. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the problem of types.

154. DOKSUM, K.A. (May 1988). On the correspondence between models in binary regression analysis and survival analysis.

155. LEHMANN, E.L. (May 1988). Jerzy Neyman, 1894-1981.

156. ALDOUS, D.J. (May 1988). Stein's method in a two-dimensional coverage problem.

157. FAN, J. (June 1988). On the optimal rates of convergence for nonparametric deconvolution problem.

158. DABROWSKA, D. (June 1988). Signed-rank tests for censored matched pairs.

159. BERAN, R.J. and MILLAR, P.W. (June 1988). Multivariate symmetry models.

160. BERAN, R.J. and MILLAR, P.W. (June 1988). Tests of fit for logistic models.

161. BREIMAN, L. and PETERS, S. (June 1988). Comparing automatic bivariate smoothers (A public service enterprise).

162. FAN, J. (June 1988). Optimal global rates of convergence for nonparametric deconvolution problem.

163. DIACONIS, P. and FREEDMAN, D.A. (June 1988). A singular measure which is locally uniform.

164. BICKEL, P.J. and KRIEGER, A.M. (July 1988). Confidence bands for a distribution function using the bootstrap.

165. HESSE, C.H. (July 1988). New methods in the analysis of economic time series I.

166. FAN, JIANQING (July 1988). Nonparametric estimation of quadratic functionals in Gaussian white noise.

167. BREIMAN, L., STONE, C.J. and KOOPERBERG, C. (August 1988). Confidence bounds for extreme quantiles.

168. LE CAM, L. (August 1988). Maximum likelihood an introduction.

169. BREIMAN, L. (August 1988). Submodel selection and evaluation in regression-The conditional case and little bootstrap.

170. LE CAM, L. (September 1988). On the Prokhorov distance between the empirical process and the associated Gaussian bridge.

171. STONE, C.J. (September 1988). Large-sample inference for logspline models.

172. Adler, R.J. and EPSTEIN, R. (September 1988). Intersection local times for infinite systems of planar brownian motions and for the brownian density process.

Copies of these Reports plus the most recent additions to the Technical Report series are available from the Statistics Department technical typist in room 379 Evans Hall or may be requested by mail from:

Department of Statistics
University of California
Berkeley, California 94720

Cost: $1 per copy.