# Optimal Designs
# Using Speckman's Minimax
# Linear Estimator

by

M. Carl Spruill*

Georgia Institute of Technology and
Purdue University

Mimeograph Series #81-49

Department of Statistics
Division of Mathematical Sciences

October, 1981

*Research support was provided by NSF Grant No. MCS-8103444.

## 1. Introduction

Speckman (1979) introduced the following minimax linear estimator. Let $Y = A\theta + \varepsilon$ where $Y \in \mathbb{R}^n$, $\theta \in \Theta$, $\Theta$ is a Hilbert space, A is a bounded linear transformation from $\Theta$ into n-dimensional Euclidean space, $E(\varepsilon) = 0$, and $E(\varepsilon\varepsilon') = \sigma^2 I$. The random vector Y is observable and $\theta \in \Theta$ is unknown. The bounded linear mapping $T: \Theta \to \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space, is given as is the linear functional $\tau \in \Theta^*$ and the scalar $\alpha$. Find the linear estimator $\ell_0'Y$ satisfying

$$\min_{\ell \in \mathbb{R}^n} \sup_{||T\theta||_{\mathcal{H}} \leq \alpha} E_\theta(\ell'Y - \tau(\theta))^2 =$$

$$\sup_{||T\theta||_{\mathcal{H}} \leq \alpha} E_\theta(\ell_0'Y - \tau(\theta))^2.$$

An example consisely represents the motivation behind the consideration of such linear estimators. Suppose that the ith component of the vector Y is $Y_i = \theta(x_i) + \varepsilon_i$, where $\{x_1,\ldots,x_n\} \subset [0,1]$ and $\theta$, the mean function, is in $W_2^2[0,2]$. Suppose $\tau(\theta) = \theta(2)$ and $T\theta = \theta''$. On the null space of T, $\theta(x) = \alpha + \beta x$. The linear estimator $\ell_0'Y$ provides robustness against possible "non-parametric" deviation from assumed linearity of the model. In this instance, as Speckman shows, the minimax linear estimator of $\theta(2)$ is the value at 2 of a data determined smoothing spline $\tilde{\theta}$ when the x values $x_1,\ldots,x_n$ are all distinct.

Our interest below is focused primarily on the problem of optimal design of experiments when Speckman's estimator is employed, although we do extend Speckman's estimator to the case of observations which are second order processes. The extension proceeds exactly along the lines of Speckman's proof. An expression for the performance is developed which enables the straightforward extension of some existing theorems on optimal designs to the present case. These optimal designs, unlike those for unbiased estimators depend upon the number N of observations. An inequality is developed which describes the relative error introduced in going from an optimal approximate theory design to a discrete design via an algorithm given by Fedorov (1972). Several example problems are solved employing a new theorem which provides sufficient conditions for optimality.

One possible classification of the optimal designs which we investigate is given by the term robust. There are many papers on this topic. See, for example Huber (1975), Kiefer (1980), Li (1981), Li and Notz (1980), Marcus and Sacks (1976), Notz (1980), and Pesotchinsky (1980). The assumed form of deviations from the model and the estimators all vary from paper to paper. None of the papers mentioned have all assumptions the same as the ones we make here. However Huber's (1975) minimax extrapolation problem is sufficiently close and we are able to make some reasonable comparisons. These may be found in example 2.

## 2. The estimator and its mean square error.

Let X be an arbitrary set and suppose that for each finite subset $\{x_{11},\ldots,x_{mn_m}\}$ of points in X one may observe the uncorrelated stochastic processes $\{Y(x_{11},t),\ldots,Y(x_{m,n_m}t): t \in T\}$ where

$$Y(x,t) = m_x(\theta,t) + \varepsilon(t)$$

$E[\varepsilon(t)] = 0$, and $K(s,t) = E[\varepsilon(s)\varepsilon(t)]$ is known. The unknown parameter $\theta$ is an element of the Hilbert space $\Theta$ and all of the mappings $\{m_x\}_{x\in X}$ are bounded and linear from $\Theta$ into the reproducing kernel Hilbert space H(K) generated by

K. Denote by B the covariance function generated by the vector of stochastic processes. That is, B is defined on the set $\Gamma \times \Gamma$, where $\Gamma = \{x_{11}, x_{12}, \ldots, x_{1n_1}, \ldots, x_{m1}, \ldots x_{mn_m}\} \times T$, by

$$B(\gamma_i, \gamma_j) = \begin{cases} K(t_i, t_j) & \text{if } \gamma_i = (x_{\ell_i u_i}, t_i) \\ & \text{and } \gamma_j = (x_{\ell_i u_i}, t_j) \\ 0 & \text{otherwise.} \end{cases}$$

The reproducing kernel Hilbert space generated by B is denoted by H(B). For the element $g \in H(B)$ the linear estimator $\langle Z, g \rangle_B$, where $Z = (Y(x_{11}, t), \ldots, Y(x_{mn_m}, t))$ has mean square error

$$V(\tau, \theta, g) = E_\theta[\langle Z, g \rangle_B - \tau(\theta)]^2$$

at $\theta \in \Theta$ for estimating the value of the linear functional $\tau$. Since

$E[\langle Z, g \rangle_B] = (m(\theta), g)_B$ and $Var(\langle Z, g \rangle_B) = ||g||_B^2$

we have

$$V(\tau, \theta, g) = ||g||_B^2 + ((m(\theta), g)_B - \tau(\theta))^2 \tag{2.1}$$

where $m: \Theta \to H(B)$. We are given the bounded linear mapping T from $\Theta$ to a Hilbert space $\mathcal{H}$ and wish to find if such exists an element $g_0 \in H(B)$ minimizing

$$\sup_{||T\theta|| \leq 1} E_\theta(\langle Y, g \rangle_B - \tau(\theta))^2.$$

Denote by M the mapping $m^*m + T^*T$.

<u>Theorem 2.1</u>. If $\mathcal{R}(M)$ is closed then whenever $\tau \in \mathcal{R}(M)$

$$(2.2)\qquad \inf_{g\in H(B)} \sup_{||T\theta||\leq 1} V(\tau,\theta,g) = \tau M\#\tau$$

where M# is the Moore-Penrose inverse relative to the ordinary orthogonal projections. If $\tau \notin \mathcal{R}(M)$ then the expression on the l.h.s. of (2.2) is $+\infty$. If $\tau \in \mathcal{R}(M)$ and $\tilde{\theta}$ is any solution to the equation $M\tilde{\theta} = \tau$ then $m(\tilde{\theta}) = g_0$ yields the unique minimax linear estimator of $\tau(\theta)$.

Proof: First suppose $\tau \notin \mathcal{R}(M)$. Since $\mathcal{R}(M)$ is closed there is a point $\varphi_0 \in \mathcal{R}^{\perp}(M) = \mathcal{N}(M)$ such that $(\varphi_0,\tau) \neq 0$. Thus from (2.1), for any $g \in H(B)$,

$$V(\tau,\alpha\varphi_0,g) = ||g||_B^2 + \alpha^2(\tau,\varphi_0)^2$$

and

$$\sup_{||T\theta||\leq 1} V(\tau,\theta,g) \geq ||g||_B^2 + \alpha^2(\tau,\varphi_0)^2$$

for all $\alpha$.

Now suppose $\tau \in \mathcal{R}(M)$. Set $\tau = M\tilde{\theta}$. First we consider the case with $||T\tilde{\theta}|| \neq 0$. In this case set $\theta_0 = \tilde{\theta}||T\tilde{\theta}||^{-1}$ and minimize

$$(2.3)\qquad V(\tau,\theta_0,g) = ||g||^2 + ((m(\theta_0),g)_B - \tau(\theta_0))^2$$

over all g. By seeking a point $g_0$ at which all directional derivatives vanish we are led to

$$(2.4)\qquad g_0 = \frac{\tau(\theta_0)}{1+||m(\theta_0)||^2}\, m(\theta_0) = m(\tilde{\theta})$$

which is easily verified to minimize $V(\tau,\theta_0,g)$ and yields $V(\tau,\theta_0,g_0) = ||m(\tilde{\theta})||^2 + ||T\tilde{\theta}||^2$. Suppose that $||T\theta|| \leq 1$. Then

$$V(\tau,\theta,g_0) = ||g_0||^2 + ((m(\theta),g_0)_B - \tau(\theta))^2$$

$$= ||g_0||^2 + [(m(\theta),g_0)_B - (m(\tilde{\theta}),m(\theta))_B - (T\tilde{\theta},T\theta)_{\mathcal{H}}]^2$$

$$= ||m(\tilde{\theta})||_B^2 + (T\tilde{\theta},T\theta)_{\mathcal{H}}^2 \leq ||m(\tilde{\theta})||_B^2 + ||T\tilde{\theta}||_{\mathcal{H}}^2.$$

Therefore, for all $g \in H(B)$ and $||T\theta||_{\mathcal{H}} \leq 1$

$$V(\tau,\theta,g_0) \leq V(\tau,\theta_0,g_0) \leq V(\tau,\theta_0,g).$$

If $\tau \in \mathcal{R}(M)$ then

$$\inf_g \sup_{||T\theta||\leq 1} V(\tau,\theta,g) \leq \sup_{||T\theta||\leq 1} V(\tau,\theta,g_0) \leq V(\tau,\theta_0,g_0)$$

$$\leq \inf_g V(\tau,\theta_0,g) \leq \sup_{||T\theta||\leq 1} \inf_g V(\tau,\theta,g).$$

Since the opposite inequality always holds we conclude that

$$\inf_g \sup_{||T\theta||\leq 1} V(\tau,\theta,g) = V(\tau,\theta_0,g_0)$$

Next, if $T\tilde{\theta} = 0$ observe that $V(\tau,\theta,g) = ||g||_B^2 + [(m(\theta),g)_B - (m(\theta),m(\tilde{\theta}))_B]^2$. Thus unless $(g - m(\tilde{\theta}), m(\theta))_B = 0$ for all $\theta \in \mathcal{N}(T)$ $\sup_{||T\theta||\leq 1} V(\tau,\theta,g) = +\infty$. Let $\mathcal{M}_T = \{m(\theta): \theta \in \mathcal{N}(T)\}$. Under the assumption that $\mathcal{R}(M)$ is closed it is readily demonstrated that $\mathcal{M}_T$ is closed. We have shown that $g - m(\tilde{\theta}) \in \mathcal{M}_T^{\perp}$ On the other hand $g = g_{\mathcal{M}_T} + g_{\mathcal{M}_T^{\perp}}$ entails $g_{\mathcal{M}_T} = m(\tilde{\theta})$ and consequently $g = m(\tilde{\theta})$. We have shown that if $T\tilde{\theta} = 0$ then $\sup_{||T\theta||\leq 1} V(\tau,\theta,g) = +\infty$ unless $g = m(\tilde{\theta})$. In that event $V(\tau,\theta,g) \overset{\theta}{\equiv} ||m(\tilde{\theta})||_B^2$. In any case, if $\tau \in \mathcal{R}(M)$ then

$$\inf_g \sup_{||T\theta||\leq 1} V(\tau,\theta,g) = \tau M\#\tau$$

since $||m(\tilde{\theta})||_B^2 + ||T\tilde{\theta}||_{\mathcal{H}}^2 = (M\tilde{\theta})\tilde{\theta} = (MM\#\tau)M\#\tau = \tau M\#\tau$. Otherwise, for all $g$ $\sup_{||T\theta||\leq 1} V(\tau,\theta,g) = +\infty$. To see that $\langle Y,g_0\rangle$ is unique it suffices to prove that if $\tau = M\tilde{\theta}_i$, $i = 1,2$, then $||m(\tilde{\theta}_1) - m(\tilde{\theta}_2)||_B = 0$. But this follows immediately from the fact that if $M(\tilde{\theta}_1 - \tilde{\theta}_2) = 0$ then $0 = (M(\tilde{\theta}_1 - \tilde{\theta}_2),(\tilde{\theta}_1 - \tilde{\theta}_2)) = ||m(\tilde{\theta}_1 - \tilde{\theta}_2)||_B^2 + ||T(\tilde{\theta}_1 - \tilde{\theta}_2)||_{\mathcal{H}}^2$. □

3. Some theorems on optimal designs.

Define the mapping $L: \Theta \to H(B) \times \mathcal{H} = V_B$ by $L\theta = (m(\theta), T\theta)$. The inner product on the Hilbert space $V_B$ is $[(u_1,u_2), (v_1,v_2)] = (u_1,v_1)_B + (u_2,v_2)_{\mathcal{H}}$. It is easily verified that $L^*L = m^*m + T^*T$. One can prove that $\mathcal{R}(L)$ is closed if and only if $\mathcal{R}(L^*L)$ is closed and that in this event $\mathcal{R}(L^*) = \mathcal{R}(L^*L)$.

Lemma 3.1. If $\mathcal{R}(M)$ is closed then for all $\tau \in \Theta$

$$\inf_g \sup_{||T\theta|| \leq 1} V(\tau,\theta,g) = \sup_{\theta \in N} \left(\frac{(\tau,\theta)^2}{||m(\theta)||^2 + ||T\theta||^2}\right) \tag{3.1}$$

where $N = \{\theta: ||m(\theta)||^2 + ||T\theta||^2 > 0\}$.

Proof: The proof proceeds as in lemma 2.3 of Spruill and Studden (1978) (except that our M here is a linear operator on $\Theta$ and $\tau$ replaces c) and will not be given here. See Nashed and Votruba (1976) Section 5.3 to verify that all the needed characteristics of the proof of lemma 2.3 are present in this case. □

In all the above the design $\{x_{11},\ldots,x_{1n_1},\ldots,x_{m,n_m}\}$ has been fixed. We wish now to examine the dependence of $\inf_g \sup_{||T\theta|| \leq 1} V(\tau,\theta,g)$ on the measure $\xi$, where $\xi$ in this case assigns masses according to $\xi(x_{ij}) = \frac{n_i}{N}$ $j = 1,\ldots,n_i$, $i = 1,\ldots,m$, and $\sum_{i=1}^{m} n_i = N$. The developments below parallel those in Spruill (1980) by identifying the mappings $m_\xi$ therein with the mappings $L_\xi: \Theta \to H(K)^{S(\xi)} \times \mathcal{H} = V_\xi$ defined by $(L_\xi\theta)(x) = (m_x(\theta), T\theta)$ for $x \in S(\xi)$. The probability measure $\xi$ is in the collection $\Xi$ of finitely supported measures on X. The inner product on $V_\xi = H(K)^{S(\xi)} \times \mathcal{H}$ is given by

$[w_1,w_2]_\xi = \Sigma\ \xi(x)[(u_{1x},v_{1x}),\ (u_{2x},v_{2x})]$. Define for $\tau \in \Theta$ and $\xi \in \Xi$

$$d_T(\tau,\xi) = \sup_{\theta \in N} \frac{(\tau,\theta)^2}{\int ||L_x\theta||^2 d\xi(x)}$$

where $N = \{\theta: \int ||L_x\theta||^2 d\xi(x) > 0\}$. We shall say that the design $\xi_0 \in \Xi$ is optimal for $\tau$ if

$$v_0 = \inf_{\Xi} d(\tau,\xi) = d(\tau,\xi_0).$$

Note that as can be seen from (3.1) and (3.2) the $\xi_0$ computed as optimal will depend upon the sample size N for in (3.1) we have $\inf_g \sup_{||T\theta||\leq 1} = N^{-1} d_{N^{-\frac{1}{2}}T}(\tau,\xi)$.

We shall concern ourselves later with this dependence. For now we take T as given and investigate the minimization of $d_T(\tau,\xi)$ over $\xi \in \Xi$.

Define

$$\mathfrak{R} = \{\int L_x^*\phi(x)d\xi(x):\ \phi \in \mathfrak{F},\ \xi \in \Xi\}$$

where $\mathfrak{F} = \{\phi: X \to H(K) \times \mathcal{H} |\ ||\phi(x)|| \leq 1\}$. We shall have occasion to invoke some or all of the following assumptions.

A1) $\Theta$ is a Hilbert space.

A2) The mappings $L_x: \Theta \to V$ are all bounded linear mappings and $\mathfrak{R}(L_\xi)$ is closed in $V_\xi$ for each $\xi \in \Xi$, where $(L_\xi\theta)(x,t) = (m_x(\theta)(t),\ T\theta)$ for $x \in S(\xi)$.

A3) There is a proper closed supporting hyperplane to $\mathfrak{R}$ at each of its boundary points.

A4) For each $\theta \in \Theta$, $\theta \neq 0$, $\sup_X ||L_x\theta|| > 0$.

<u>Lemma 3.2</u>. Let A1-A4 hold. Then

a) $\beta\ \tau \in \mathfrak{R}$ implies $v_0 \leq \frac{1}{\beta^2}$, and

b) $\beta\ \tau \in \partial\mathfrak{R}$ implies $v_0 \geq \frac{1}{\beta^2}$.

Proof: This lemma corresponds exactly to Lemma 3.1 of Spruill (1980) with the replacements we have indicated. The proof proceeds in a straightforward manner by making those replacements so we will not give all of it here. However the proof of part (b) of Spruill (1980) contains an error which we here take the opportunity to rectify. The error occurs following inequality (14) therein; $(\beta-\varepsilon)\tau$ may not be in $\mathcal{R}$ for all $\varepsilon$ sufficiently small. We proceed as follows. Since $\beta\,\tau \in \partial\mathcal{R}$ there is a sequence of points $\{r_n\}_{n\geq 1} \subset \mathcal{R}$ such that $||r_n - \beta\tau||_\Theta \to 0$. Since $r_n(\theta) = \sum_{j=1}^{m_n} \alpha_{jn}(L_{x_{jn}}(\theta), \phi_{jn}) \leq \sup_X ||L_x(\theta)||$ for all n we conclude that $\beta(\tau,\theta) \leq \sup_X ||L_x(\theta)||$. Therefore $\sup_X ||L_x(\theta)|| = \beta\tau(\theta)$ which corresponds to equation (16). □

Theorem 3.1. Under conditions A1)-A4) if $\tau \in \mathcal{R}(L_\xi^* L_\xi)$ for some design $\xi \in \Xi$ then $d(\tau,\xi_0) = v_0$ and $\xi_0 \in \Xi$ if and only if there is a function $\phi \in \mathfrak{F}$ such that $||\phi||_v \doteq 1$ and $\int L_x^* \phi(x) d\xi_0(x)$ is

i) proportional to $\tau$ and

ii) in $\mathcal{R} \cap \partial\mathcal{R}$.

Proof: As above when the proper replacements are made the same arguments go through. Note that, as is not made clear in Theorem 3.1 of Spruill (1980) the G-inverse $M^\#$ is relative to the orthogonal projectors. □

Theorem 3.2 of Spruill also goes through. However we have a much improved version which we state below in its stead. Parts of the improved portion of the theorem have already been proven in Spruill (1981) and we refer the reader to that proof, again with the obvious identifications. Since the proof of the remaining portion was not given in Spruill (1980) we record it below. Let

$\Delta = \{\theta \in \Theta: (\tau,\theta) = 1\}$ and recall that $L_x = (m_x(\theta), T_\theta)$, an ordered pair in $V = H(K) \times \mathcal{H}$ for all $x \in X$.

Theorem 3.2. Suppose there is a point $\delta_0 \in \Delta$ and a design $\xi_0 \in \Xi$ satisfying

i) $S(\xi_0) \subset \{x: ||L_x(\delta_0)||_V = \sup_X ||L_x(\delta_0)||_V\}$

ii) $\int L_x^* L_x \ \delta_0 \ d\xi_0(x) = \alpha\tau$ for some $\alpha > 0$, and

iii) $\int ||L_x\theta||_V^2 \ d\xi_0(x) = 0$ entails $\tau(\theta) = 0$.

Then $\xi_0$ satisfies $d_T(\tau,\xi_0) = \inf_\Xi d_T(\tau,\xi)$ and

iv) $\inf_\Delta \sup_X ||L_x(\delta)||_V^2 = \sup_X ||L_x(\delta_0)||_V^2$.

The conditions on $m_x$, T, and $\tau$ in order for this to hold are $m_x$ all bounded and linear, T bounded and linear and $\tau \in \Theta$. Conversely, if conditions A1)-A4) also hold and an optimal design $(d_T(\tau,\xi_0) = \inf_\Xi d_T(\tau,\xi) < \infty)$ $\xi_0 \in \Xi$ exists then a point $\delta_0 \in \Delta$ may be found satisfying all conditions i) through iv).

Proof: The proof of the first part as given in Spruill (1981) goes through upon noting that $\Delta \cap N(\xi_0) \cup \Delta \cap N^c(\xi_0) = \Delta$ (called U there) and $\Delta \cap N^c(\xi_0) = \phi$ by iii). Now suppose that A1)-A4) hold and that $\xi_0 \in \Xi$ is optimal. By theorem 3.1 there is a function $\phi: X \to H(K) \times \mathcal{H}$ such that $||\phi(x)|| \equiv 1$,

(3.3) $$\int L_x^* \ \phi(x) d\xi_0(x) = \beta\tau,$$

and $\beta\tau \in \partial\mathcal{R}$. By A3 there is a $\lambda \neq 0$, $\lambda \in \Theta$ such that $(\lambda,r) \leq \beta(\lambda,\tau)$ for all $r \in \mathcal{R}$. Since by A4 $\sup_X ||L_x(\lambda)|| > 0$ we may find a sequence of points $\{x_n\}$ in X satisfying $||L_{x_n}(\lambda)|| \uparrow \sup_X ||L_x(\lambda)||$ and $||L_{x_n}(\lambda)|| > 0$. Set $r_n = L_{x_n}^* \dfrac{L_{x_n}(\lambda)}{||L_{x_n}(\lambda)||}$. Then $r_n \in \mathcal{R}$ for all n and since (3.3) holds

$$\lim_{n\to\infty} ||L_{x_n}(\lambda)|| \le \sum_{i=1}^{m} \xi_0(y_i)(L^*_{y_i}\ \phi(y_i),\ \lambda)$$

$$= \sum_{i=1}^{m} \xi_0(y_i)(\phi(y_i),\ L_{y_i}\lambda)$$

$$\le \sup_X ||L_x(\lambda)||$$

with strict inequality unless $||L_x(\lambda)|| \equiv \sup_X ||L_x\lambda||$ for $x \in S(\xi_0)$. Set $\delta_0 = \frac{\lambda}{\tau(\lambda)}$ ($\tau(\lambda) \neq 0$ since $\alpha\tau(\lambda) > 0$). Clearly i) is satisfied. From above we also conclude that for all $x \in S(\xi_0), \phi(x) = k_x L_x(\lambda)$. This in turn implies, for $x \in S(\xi)$ that $\phi(x) = \frac{L_x(\lambda)}{||L_x(\lambda)||}$. Therefore

$$\int L^*_x\ \phi(x)d\xi_0(x) = [\int L^*_x L_x(\delta_0)d\xi_0(x)][\sup_X ||L_x(\lambda)||]^{-1}$$

and we see that ii) is also satisfied. If iii) is not satisfied then there is a sequence $\theta_n$ such that $\int ||L_x(\theta_n)||^2 d\xi_0(x) \to 0$ and $\tau(\theta_n) \to t \neq 0$. This implies $d(\tau,\xi_0) = +\infty$ which contradicts our assumptions. We conclude that iii) is satisfied and consequently that (iv) also is satisfied. □

When $\Theta$ is finite dimensional, X is compact, and the mappings $L_x\theta$ are continuous in x for each fixed $\theta$ then one may prove much stronger theorems than 3.1 and 3.2. We state, without proof, one such theorem. We shall need the following conditions.

B1) The mappings $L_x$: $\Theta \to V$ are linear for each $x \in X$.

B2) There is a topology on X for which X is compact and one point subsets are Borel measurable.

B3) For each fixed $\theta \in \Theta$ the mappings $L_x\theta$: $X \to V$ are continuous in X.

B4) $\Theta$ is a finite dimensional Hilbert space.

Theorem 3.3. Under conditions A4) and B1)-B4) there is an optimal design $\xi_0 \in \Xi$ for estimating $\tau(\theta)$ whose support contains no more than dim $\Theta$ points. In addition conditions A1)-A4) are satisfied so that a point $\delta_0$ exists in $\Delta = \{\theta: (\tau,\theta) = 1\}$ satisfying i)-iv) of Theorem 3.2.

Finally one can prove the following theorem, which is the analogue of Theorem 4.1 in Spruill (1980), using the same techniques as employed there.

Theorem 3.4. If $L_x$ is a bounded linear operator for each $x \in X$, there is a constant $k > 0$ such that for all $\theta$

$$\sup_X ||L_x\theta||_V \geq k||\theta||,$$

and (A1) holds, then (A3) holds.

## 4. Finding good exact designs.

We shall demonstrate the effectiveness of the theorems above in producing designs which minimize, for a given T and $\tau$, the function $d_T(\xi,\tau)$. However we should recall that for a given N, T, and $\tau$ we really would like to minimize the expression in (3.1) which may be written as $N^{-1}d_{T/\sqrt{N}}(\tau,\xi)$. In the usual design problem using the best linear unbiased estimator the operator T is the zero operator. In that case one may employ the optimal approximate theory design to pass, for each N, to a good design $\xi$ in

$\Xi_N = \{\xi \in \Xi$: # supp$(\xi) \leq N$ and $\exists\, M \leq N$ with $M\xi(x)$ an integer $\forall x \in \text{supp}(\xi)\}$

Fedorov (1972) gives such a procedure and inequalities which provide a measure of the departure from optimality of the design so constructed. In our present case we observe the possibility that the designs which minimize $d_{T/\sqrt{N}}(\tau,\xi)$ will depend upon N. Thus a change in the routine of passing from the approximate theory to the exact theory has been introduced. Moreover, since the assumptions employed by Fedorov in providing the inequalities do not hold it is not clear what procedure should be employed in the construction of a good exact design from the optimal approximate theory design. We shall prove in this section that Fedorov's procedure continues in the present case to provide as good a method of finding a good exact design from the approximate as it does when unbiased linear estimators are employed.

There is a reasonable alternative to the modus operandi upon which we have embarked which bears mentioning. If in the definition of the linear estimator we measure the allowable deviation from the model relative to the size of the sample we may circumvent some ensuing difficulties. Specifically, if we take the supremum over $\theta$ satisfying $||T\theta|| \leq N^{-\frac{1}{2}}$ the minimax mean square error is $N^{-1}d_T(\tau,\xi)$. For this estimator the approximate theory optimal design (if one exists) will not depend upon N. Still the assumptions employed by Fedorov in bounding the error in passing from the optimal approximate theory design to a good design in $\Xi_N$ do not hold for the functional $d_T(\tau,\xi)$.

In the following $A = A_0^*A_0$, $B = B_0^*B_0$, and $D = D_0^*D_0$ where $A_0$, $B_0$, and $D_0$ are all bounded linear operators from $\Theta$ into Hilbert spaces and all their ranges are closed. For $\tau \in \Theta$ fixed define

$$\tilde{L}_D(A) = \left[\sup_{\theta \in N(A,D)} \frac{(\tau,\theta)^2}{\theta(A+D)\theta}\right]^{-1}$$

where we take $[+\infty]^{-1} = 0$ and $N(A,D) = \{\theta\colon \theta(A+D)\theta > 0\}$.

Lemma 4.1. For all scalars $k \geq 1$

$$k\tilde{L}_D(A) \geq \tilde{L}_D(kA) \quad \text{and} \quad k\tilde{L}_D(A) \geq \tilde{L}_{kD}(A).$$

Proof: The same proof works for both. We show the first. Since $\theta(kA+D)\theta \geq \theta(A+D)\theta$ we have $N(A,D) \subset N(kA,D)$. Also for all $\theta \in N(A,D)$ we have $\theta(kA+kD)\theta \geq \theta(kA+D)\theta$ so that

$$k\tilde{L}_D(A) \geq \left[\sup_{\theta \in N(A,D)} \frac{(\tau,\theta)^2}{\theta(kA+D)\theta}\right]^{-1}$$

$$\left[\sup_{\theta \in N(kA,D)} \frac{(\tau,\theta)^2}{\theta(kA+D)\theta}\right]^{-1} = L_D(kA). \quad \square$$

Lemma 4.2. $\tilde{L}_D(A+B) \geq \tilde{L}_D(A)$.

Proof: We have $N(A,D) \subset N(A+B,D)$. Also $\theta(A+B+D)\theta \geq \theta(A+D)\theta$ so

$$\tilde{L}_D(A) \leq \left[\sup_{\theta \in N(A,D)} \frac{(\tau,\theta)^2}{\theta(A+B+D)\theta}\right]^{-1}.$$

We know that if $\tau \notin \mathcal{R}(A+D)$ then $\tilde{L}_D(A) = 0$ so that $\tilde{L}_D(A+B) \geq \tilde{L}_D(A)$ is trivially true. If $\tau \in \mathcal{R}(A+D)$ then for $\theta \in N(A+B,D) - N(A,D)$, $(A+D)\theta = 0$. Since $\mathcal{R}(A+D) = \mathcal{N}^{\perp}((A+D)^*) = \mathcal{N}^{\perp}(A+D)$, $(\tau,\theta) \equiv 0$ for all such $\theta$. Therefore, on $N(A+B,D) - N(A,D)$, $\frac{(\tau,\theta)^2}{\theta(A+B+D)\theta} \equiv 0$ and we conclude that

$$\sup_{\theta \in N(A+B,D)} \frac{(\tau,\theta)^2}{\theta(A+B+D)\theta} = \sup_{\theta \in N(A,D)} \frac{(\tau,\theta)^2}{\theta(A+B+D)\theta}.$$

This proves that $\tilde{L}_D(A) \leq \tilde{L}_D(A+B)$. $\square$

Let N be a fixed positive integer and T, as above, be a bounded linear operator from $\Theta$ into $\mathcal{H}$. Suppose that $\xi^* \in \Xi$ satisfies

$$\inf_{\xi \in \Xi} d_{T/\sqrt{N}}(\tau,\xi) = d_{T/\sqrt{N}}(\tau,\xi^*)$$

and that $\#\,\mathrm{supp}(\xi^*) = r < N$. Also suppose that $\xi_N^* \in \Xi_N$ satisfies

$$\inf_{\xi \in \Xi_N} d_{T/\sqrt{N}}(\tau,\xi) = d_{T/\sqrt{N}}(\tau,\xi_N^*).$$

Define the measure $\tilde{\xi}_N \in \Xi_N$ from $\xi^*$ as follows. Denoting by [ ] the least integer function defined by [x] = smallest integer greater than or equal to x, assign $[(N-r)\xi^*(x_i)]$ observations at $x_i \in \text{supp}(\xi^*)$. Since $[(N-r)\xi(x_i)] < (N-r)\xi(x_i) + 1$ this uses $\leq \sum_{i=1}^{r} ((N-r)\xi(x_i) + 1 = N - r + r = N$ observations. Assign the other observations in any manner and denote the proportion of the total N at $x_i$ by $\tilde{\xi}_N(x_i)$.

Theorem 4.1. The design $\tilde{\xi}_N$ constructed from $\tilde{\xi}^*$ as above satisfies

$$0 \leq 1 - \frac{d_{T/\sqrt{N}}(\tau, \xi_N^*)}{d_{T/\sqrt{N}}(\tau, \tilde{\xi}_N)} \leq \frac{r}{N-r}$$

and

$$0 \leq 1 - \frac{d_{T/\sqrt{N}}(\tau, \xi^*)}{d_{T/\sqrt{N}}(\tau, \xi_N^*)} \leq \frac{r}{N}$$

whenever $M(\xi) = \int L_X^* L_X d\xi(x)$ has closed range for all $\xi \in \Xi$.

Proof: Both relationships above are a consequence of

$$(4.1) \qquad \tilde{L}_D[M(\xi^*)] \geq \tilde{L}_D[M(\xi_N^*)] \geq \tilde{L}_0[M(\xi_N)] \geq (1-\tfrac{r}{N})\tilde{L}_0[M(\xi^*)]$$

where $D = T^*T/_N$ so we show (4.1). The left-most inequality follows from $\Xi \supset \Xi_N$. The middle inequality follows from the fact that $\tilde{\xi}_N \in \Xi_N$ by definition of $\xi_N^*$. The proof of the right-most inequality uses the facts established in lemmas 4.1 and 4.2. Since $\frac{N}{N-r} > 1$

$$\frac{N}{N-r} \tilde{L}_D[M(\tilde{\xi}_N)] \geq \tilde{L}_D(\frac{NM(\tilde{\xi}_N)}{N-r})$$

$$= \tilde{L}_D[\frac{\sum_{i=1}^{n} ([(N-r)\xi^*(x_i)] + \alpha_i)L_{x_i}^* L_{x_i}}{N-r}],$$

where $N\tilde{\xi}_N(x_i) = [(N-r)\xi^*(x_i)] + \alpha_i$. Since $\frac{[(N-r)\xi^*(x_i)]}{N-r} \geq \frac{(N-r)\xi^*(x_i)}{N-r}$ we have

$$\frac{N}{N-r}\,\tilde{L}_D[M(\tilde{\xi}_N)] \geq \tilde{L}_D[M(\xi^*) + \sum_{i=1}^{\tau} \gamma_i L^*_{x_i} L_{x_i}]$$

where $\gamma_i = \dfrac{[(N-r)\xi^*(x_i)] + \alpha_i}{N-r} - \xi^*(x_i) \geq \dfrac{\alpha_i}{N-r} \geq 0$. By Lemma 4.2 $\tilde{L}_D[M(\xi^*) + \Sigma\gamma_i L^*_{x_i} L_{x_i}] \geq \tilde{L}_D[M(\xi^*)]$ concluding the proof of the theorem. □

Note that r itself may conceivably depend upon N. We do not investigate this general question here. In the examples below r remains bounded for all N so that by making N sufficiently large the relative loss which comes about by using $\tilde{\xi}_N$ instead of $\xi^*_N$ can be made as small as desired.

## 5. Some examples.

In this section we investigate the application of the theorems to some examples of varying complexity. In each example the strategy is

a) to find if possible a point $\delta_0 \in \Delta = \{\theta\colon (\tau,\theta) = 1\}$ satisfying

$$\inf_{\Delta} \sup_{X} ||L_x\theta||^2 = \sup_{X} ||L_x\delta_0||^2, \text{ and}$$

b) to find a measure $\xi \in \Xi$ satisfying

$$\int (L_x\delta_0, L_x\theta)d\xi(x) \overset{\theta}{\equiv} \sup_{X} ||L_x\delta_0||^2\tau(\theta)$$

$$\text{and } S(\xi) \subset \{x\colon ||L_x(\delta_0)||^2 = \sup_{X} ||L_x\delta_0||^2\}.$$

Without checking the assumptions A1)-A4) we can not be sure that this procedure will yield an optimal design. However, even if A1)-A4) are not satisfied it may work and even if A1)-A4) are satisfied one must identify, in using Theorem 3.2, the $\delta_0$ in a) and the design in b).

Example 1. We suppose that we may observe, for $x \in [a,b]$, $a < b$,

$$Y(x) = \theta_0 + \theta_1 x + \varepsilon$$

where $E(\varepsilon) = 0$, $E(\varepsilon^2) = 1$, and $\theta_0$ and $\theta_1$ are unknown. We wish to estimate the value $\theta_0 + \theta_1 c$ using an estimator which guards against $|\theta_1|$ too large. The estimator we employ is Speckman's minimax linear estimator with $\Theta = \mathbb{R}^2$ and $T\theta = \frac{\theta_1}{\alpha}$, where $\alpha > 0$ is a preassigned constant. Thus we shall be finding the proper proportions of observations to be taken at points which we shall determine in [a,b] in order to minimize the maximum mean square error over the set $\{\theta: |\theta_1| \leq \alpha\}$. As we have seen above, we must prescribe an N in order to proceed. Having done this we shall prove below that if $\alpha^2 N > z_0$ then the optimal design takes all measurements at a and b, the optimal proportions being respectively

(5.1) $$\xi^*(a) = \frac{1}{2} - \frac{1}{b-a} J\bar{\theta}_1 \quad \text{and} \quad \xi^*(b) = 1 - \xi^*(a),$$

where $z_0 = 2[(c-b)(b-a)]^{-1}$, $J = \frac{(b-a)^2}{4} + \frac{1}{N\alpha^2}$, and $\bar{\theta}_1 = [c - (\frac{b+a}{2})]^{-1}$. If $\alpha^2 N < z_0$ then the optimal design turns out to place all of its mass at the point b. This may seem counter-intuitive since $\theta_0 + \theta_1 c$ would not be linearly estimable. But recall that the estimator we are utilizing is not unbiased. We begin our demonstration with step a). In our present example $H(K) = \mathbb{R}^1 = \mathcal{H}$, $L_x\theta = (\theta_0 + \theta_1 x, \frac{\theta_1}{\sqrt{N\alpha}}) \in \mathbb{R}^2$. Since $\Delta = \{\theta \in \mathbb{R}^2: \theta_0 + c\theta_1 = 1\}$ we seek $\theta_1$ which minimizes

$$\sup_{a \leq x \leq b} ||L_x\theta||^2 = \sup_{a \leq x \leq b} ((\theta_0 + \theta_1 x)^2 + \frac{\theta_1^2}{\alpha^2 N})$$

$$= \max\{((1-\theta_1)(c-a))^2, (1-\theta_i(c-b))^2\} + \frac{\theta_1^2}{\alpha^2 N} .$$

A graphical representation of the three functions $f_1(t) = (1 - t(c-a))^2$, $f_2(t) = (1-t(c-b))^2$, and $f_3(t) = {t^2}/{N\alpha^2}$ is helpful in solving this problem. A rough sketch is given in Figure 1.

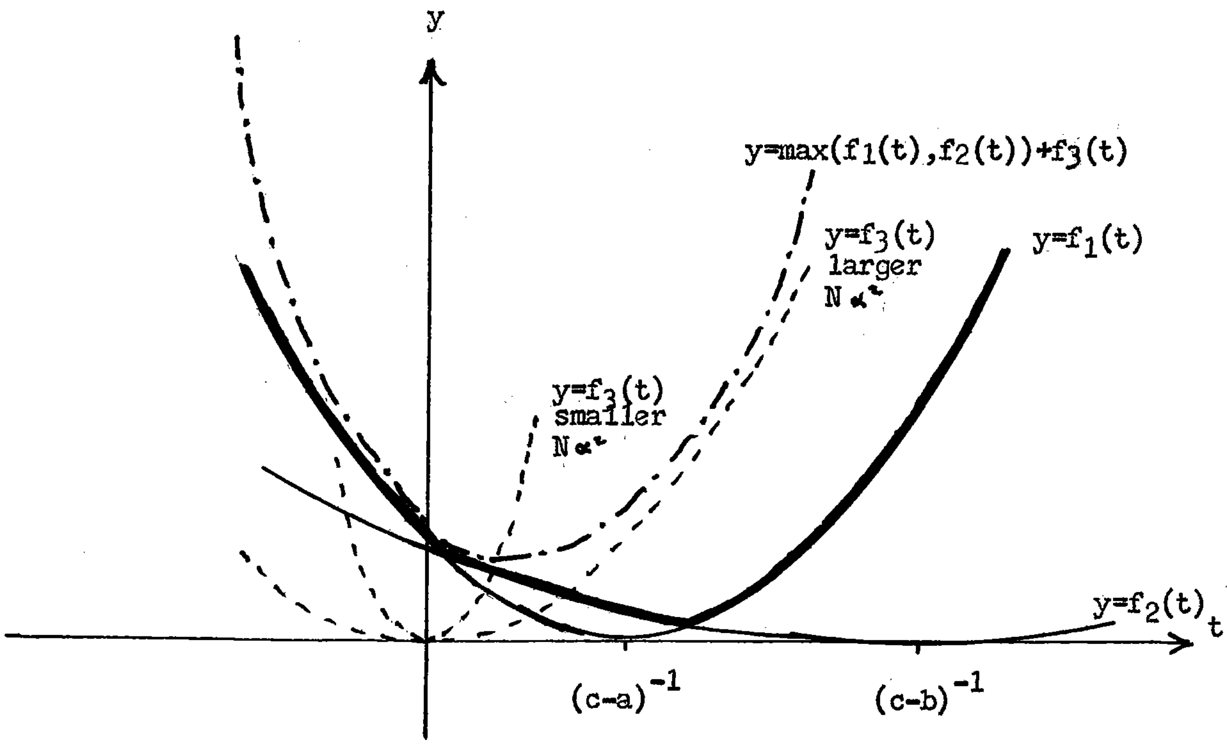


FIG.1

It is clear from this rough sketch that for $\alpha^2 N$ sufficiently large the minimizing value of $\theta$ will be $\bar{\theta}_1 = [c - (\frac{a+b}{2})]^{-1}$, the point at which the two parabolas intersect. Just how large $\alpha^2 N$ must be is determined by the fact that for $\alpha^2 N$ small the minimum will occur for $\bar{\bar{\theta}}_1 \in (0,\bar{\theta}_1)$ minimizing $f_2(\theta_1) + f_3(\theta_1)$. This point is $\bar{\bar{\theta}}_0 = \dfrac{c-b}{(c-b)^2 + \frac{1}{N^{\alpha 2}}}$ and the minimum value is

$1/_{N\alpha^2}\,[(c-b)^2 + \frac{1}{\alpha^2 N}]^{-1}$. As $N\alpha^2$ is taken larger $\bar{\bar{\theta}}_1$ will approach $\bar{\theta}_1$ and will equal it for $N\alpha^2 = 2[(c-b)(b-a)]^{-1}$. For larger values of $N\alpha^2$ the minimum will continue to occur at $\bar{\theta}_1$ and the minimum value there will be $J\bar{\theta}_1^2$. Now we seek the designs via part b). There are two cases.

Case I. $N\alpha^2 \geq z_0$. Since the supremum over $x \in [a,b]$ of $||L_x\bar{\theta}_1||^2$ occurs at a and b only the support of any optimal design must be contained therein. We seek therefore $\xi^* \in [0,1]$ satisfying

$$\xi^*(\theta_0 + \theta_1 a)(\bar{\theta}_0 + \bar{\theta}_1 a) + (1 - \xi^*)(\theta_0 + \theta_1 b)(\bar{\theta}_0 + \theta_1 b)$$

$$+ \frac{\theta_1\bar{\theta}_1}{N\alpha^2} \equiv (\theta_0 + \theta_1 c)J\bar{\theta}_1^2 \quad \text{for all } \theta_0 \text{ and } \theta_1.$$

Writing $\bar{\theta}_0 + \bar{\theta}_{1a} = (\frac{a-b}{2})\bar{\theta}_1$ and $\bar{\theta}_0 + \bar{\theta}_1 b = (\frac{b-a}{2})\bar{\theta}_1$ we have $(1-2\xi^*)(\frac{b-a}{2}) = \bar{\theta}_1^2 J$. The solution $\xi^*$ is given in (5.1) and it remains only to verify that

$$\xi^* a(\bar{\theta}_0 + \bar{\theta}_1 a) + (1-\xi^*)b(\bar{\theta}_0 + \bar{\theta}_1 b) + \frac{\bar{\theta}_1}{N\alpha^2} = cJ\bar{\theta}_1^2.$$

This is easily done and we proceed to case II.

Case II. $N\alpha^2 < z_0$. In this case one may check that the unique maximum of $(1 - \bar{\bar{\theta}}_1(c-x))^2$ over $x \in [a,b]$, where $\bar{\bar{\theta}}_1 = (c-b)\gamma = (c-b)/[(c-b)^2 + \frac{1}{N\alpha^2}]$ occurs at $x = b$. Consider the measure which places all mass at b. Since $(1 - \bar{\bar{\theta}}_1(c-b))^2 + \frac{\bar{\bar{\theta}}_1}{N\alpha^2} = \gamma/_{N\alpha^2}$ it suffices to check that for all $(\theta_0,\theta_1)' \in \mathbb{R}^2$

$$(\theta_0 + b\theta_1)(\bar{\bar{\theta}}_0 + \bar{\bar{\theta}}_1 b) + \frac{\theta_1\bar{\bar{\theta}}_1}{N\alpha^2} \equiv \frac{\gamma}{N\alpha^2}(\theta_0 + \theta_1 c). \tag{5.2}$$

Since $(\bar{\bar{\theta}}_0 + \bar{\bar{\theta}}_1 b) = (1 - \bar{\bar{\theta}}_1(c-b)) = \gamma/_{N\alpha^2}$ we have shown the $\theta_0$ portion of the identity. For $\theta_1$ the left hand side of (5.2) setting $\theta_0 = 0$ and $\theta_1 = 1$ is just $\frac{b\gamma}{N\alpha^2} + \frac{(c-b)}{N\alpha^2} = \frac{c}{N\alpha^2}$ as required. Let us conclude this example by giving the explicit estimators corresponding to the two cases. Recall that for a design in $\Xi_N$ (see (2.4)) the point $g_0$ in the reproducing kernel Hilbert space satisfies $m_\xi(\tilde{\theta}) = g_0$ where $\tilde{\theta}$ is any solution to $M\tilde{\theta} = \tau$ and $M = N \int m_x^* m_x \, d\xi(x) + \frac{1}{\alpha^2} T^*T$, which in this example is just

$$N(\sum \begin{bmatrix} 1 & x \\ x & x^2 \end{bmatrix} \xi(x) + \frac{1}{N\alpha^2} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}).$$

The process of solution has provided us with $\tilde{\theta}$ at no extra cost in time and effort. Thus in the case $N\alpha^2 \geq z_0$ we have $N^{-1}M\bar{\theta} = J\bar{\theta}_1^2\tau$ so that $\tilde{\theta}_I = \frac{N^{-1}\bar{\theta}}{J\bar{\theta}_1^2}$. In the case of $N\alpha^2 < z_0$ we similarly have

$\tilde{\theta}_{II} = \frac{\alpha^2\bar{\bar{\theta}}}{\gamma}$. The reproducing kernel Hilbert space for this case is just $\mathbb{R}^1$ so

$$<Y, m(\tilde{\theta}_I)> = ( \sum_{x=a} Y(x)(1,a)\bar{\theta} + \sum_{x=b} Y(x)(1,b)\bar{\theta}) \frac{1}{NJ_1^{-2}}$$

and

$$<Y, m(\tilde{\theta}_{\pi})> = \frac{\alpha^2}{\gamma} (1,b)\bar{\bar{\theta}} \sum Y(b)$$

Example 2. Suppose that we may observe $Y(x) = \theta(x) + \varepsilon$, where $E(\varepsilon) = 0$, $E(\varepsilon^2) = 1$, and $\theta$ is an unknown function in the Sobolev space $W_2^2[a,c]$ of all real valued functions on [a,c] whose first derivative is absolutely continuous and whose second derivative is in $L_2[a,c]$. We assume that based upon observations taken in the interval [a,b], $b \in (a,c)$ fixed, we are to provide a good estimate of $\theta(c)$. It is assumed that the function $\theta$ is nearly linear, in the sense that $||\theta''||_2^2 = \int_a^c (\theta''(s))^2 ds$ is small, say $\leq \alpha^2$, and that in order to guard against such possible departures from linearity Speckman's estimator will be employed with $T = \frac{1}{\alpha}D^2$. What are the optimal experimental designs for this problem?

We proceed as in the last example. Given a value N we seek $\theta_0$ satisfying condition a) of 5 and a design $\xi^*$ satisfying b). We state the solution below. The proof that $\theta_0$ and $\xi^*$ do indeed satisfy conditions a) and b) is contained in Spruill (1981). The function $\theta_0$ in $W_2^2[a,c]$ turns out to be a spline with knots at $x_1(\alpha^2N)$ and $x_2 \equiv b$. Define

$$\eta_0 = \frac{(b-a)^2}{24}\left[\frac{1}{b-a} + \frac{1}{2(c-b)}\right]^{-1}.$$

For $\alpha^2N \leq \eta_0^{-1}$, $x_1(\alpha^2N) = a$. For $\eta_0^{-1} < \alpha^2N$, $x_1(\alpha^2N)$ is the unique real solution in [a,b] to the equation

$$\frac{\alpha^2N(b-x)^2}{24} - \frac{1}{2(c-b)} = \frac{1}{b-x}. \tag{5.3}$$

The two knots are located in [a,b] at the points at which $\theta_0$ achieves its maximum absolute value. The exact formula for $\theta_0$ is, setting $\eta = (\alpha^2N)^{-1}$,

$$\theta_0(x) = [s^2(\eta) + \eta z^2(\eta)]^{-1}[\eta z(\eta)\sum_{i=1}^{2}(-1)^i\ \phi_{x_i}(x) + \int_a^c h_x(s)h_c(s)ds]$$

where $\phi_{x_1}(x) = \frac{x-x_2}{x_1-x_2}$, $\phi_{x_2}(x) = \frac{x-x_1}{x_2-x_1}$,

$$h_x(s) = (x-s)_+ - \sum_{i=1}^{2}\phi_{x_i}(x)(x_i-s)_+,$$

$$s^2(\eta) = \int_a^c h_c^2(s)ds = \frac{(c-b)^2(c-x_1)}{3},$$

and $z(\eta) = |\phi_{x_1}(c)| + |\phi_{x_2}(c)| = 2\left(\frac{c-(\frac{a+b}{2})}{b-x_1}\right)$.

The optimal design $\xi^*$ places mass $\frac{|\phi_{x_1}(c)|}{|\phi_{x_1}(c)|+|\phi_{x_2}(c)|} = \frac{c-b}{c-b+c-x_1}$ at $x_1$ and $\frac{c-x_1}{c-b+c-x_1}$ at $x_2 = b$. If both of these are rational of the form $\frac{n_1}{N}$ and $\frac{n_2}{N}$ then the corresponding estimator is $\langle Y,\tilde{\theta}\rangle$ where $\tilde{\theta} = \gamma\theta_0$. We may determine $\gamma$ by $M\tilde{\theta} = \tau$ since we know $\int L_x^* L_x \theta_0 d\xi^*(x) = M_\xi \theta_0 = \eta[s^2(\eta) + \eta z^2(\eta)]^{-1}$ and $NM_\xi = M$. Thus $\tilde{\theta} = \frac{[s^2(\eta)+ z^2(\eta)]}{N\eta} \theta_0$. Again, for this scalar observation problem, the reproducing kernel Hilbert space is $\mathbb{R}^1$ and we have the associated estimator

$$\langle Y,\tilde{\theta}\rangle = \sum_{x=x_1} \tilde{\theta}(x_1)Y(x) + \sum_{x=x_2} \tilde{\theta}(x_2)Y(x)$$

$$= \frac{z(\eta)}{N} [\sum_{x=x_2} Y(x) - \sum_{x=x_1} Y(x)]$$

$$+ \frac{1}{N} \int_a^c (\sum_x Y(x)h_x(s))h_c(s)ds$$

Its maximum mean squared error over $||\theta''||_2 \leq \alpha$, where $\alpha > 0$, is $\frac{[s^2(\eta)+ z^2(\eta)]}{\eta N}$ where $\eta = (N\alpha^2)^{-1}$. Moreover, among all possible choices of N points the above choice yields this smallest possible value of the mean square error.

Remark. If the model is $\theta(x) = \alpha + \beta x$, $\theta(c)$ is to be estimated, best linear unbiased estimators are used, and observations are allowed only in [a,b] then the usual theory shows the optimal design to place $\frac{c-b}{c-b+c-a}$ at a and $\frac{c-a}{c-b+c-a}$ at b. Our optimal designs coincide for $\alpha^2 N$ small. In any case our designs are the usual optimal ones on the shorter interval $[x_1,b]$. Huber (1975) employed somewhat different assumptions and arrived at similar conclusions. Huber assumed that observations were available on $x \in [0,\infty)$, $\theta(-1)$ was to be estimated by a minimax linear estimator, and the contamination was $\sup_{-1 \leq x < +\infty} |\theta''(x)| \leq \varepsilon$. The same results obviously hold if the interval is

$(-\infty,0]$ and $\theta(1)$ is to be estimated by replacing his $1/\gamma$ with $-1/\gamma$ as the location of the left-most point of support in the optimal design. In this context, the right-most point is 0. As Huber points out, his optimal designs also assign weights corresponding to those of the usual optimal designs on $[-\frac{1}{\gamma}, 0]$ to extrapolate to 1. An examination of our arguments reveals that our results continue to be valid if $a = -\infty$ by setting $\frac{1}{\eta_0} = 0$. We shall compare the two optimal designs at $N\alpha^2$ values. In Huber's notation $n = N$, $\varepsilon^2 = \alpha^2$, and we take $\sigma^2 = 1$. Since the class of contaminants assumed by Huber $||\theta''||_\infty \leq \alpha$ is smaller than our class $||\theta''||_2 \leq \alpha$ on the unit interval one might conjecture that the allowable contamination $\alpha'$ in $||\cdot||$ model to place the left point $x_1 = \frac{-1}{\gamma_1}$ at the left-point $x = \frac{-1}{\gamma}$ of Huber's using $\alpha$ should be smaller. This is indeed the case for large N as we now demonstrate. Using his equation (6.22) and our equation (5.3) we see that $1 = \dfrac{12\ \gamma_1^2(1+2\gamma_1)}{8\ \gamma^2(1+2\gamma)(\gamma/1+\gamma)}$. As $N\alpha^2 \to \infty$ $\delta_1$ and $\delta$ both become large and we have $\frac{\gamma_1}{\gamma} \to (2/3)^{1/3}$. Thus $\frac{x_1}{x} \to (2/3)^{1/3}$ and $\alpha'$ is roughly $(2/3)\alpha$ for large equal sample sizes N.

In the next example we require the notion of a band limited function. The reader is referred to Slepian and Pollack (1961) for details. Denote the space of Lebesgue measurable complex valued functions $f\colon \mathbb{R}^1 \to \mathbb{C}$ satisfying $\int |f|^p < \infty$ by $L^p$. There is an isometry $F\colon L^2 \to L^2$, called the Fourier-Plancherel transform, which satisfies for all $f \in L^1 \cap L^2$, $w \in \mathbb{R}^1$

$$F(f)(w) = \int_{-\infty}^{\infty} f(t)e^{-iwt}dm(t),$$

where m is Lebesgue measure normalized by $(2\pi)^{-\frac{1}{2}}$ (see Rudin 1966)). If $a > 0$ then the space $\mathcal{B}_a$ of band limited functions is that subset of $L^2$ whose Fourier-Plancherel transforms vanish off of the interval $[-a,a]$. We denote this set by $\mathcal{B}_a$. We have, defining the operators $B_a$ by $B_a f = fI_{[-a,a]}$,

$$\mathcal{B}_a = \{f \in L^2 : f = (F^{-1}B_aF)f\}.$$

The operators $F^{-1}B_aF$ are easily verified to be bounded and linear and $\mathcal{B}_a$ are closed subspaces of $L^2$. For $f \in L^2$ denote by $f_{\mathcal{B}_a}$ the projection of f onto $\mathcal{B}_a$.

Before relating the precise assumptions pertaining to Example 3 we offer some interpretive comments. We have available for "observation" real stationary generalized stochastic processes $Y - \theta = W'$, where $\theta \in L^2$ is real valued and W' is "white noise". We wish to estimate the unknown value of $\tau(\theta) = \int_{-\infty}^{\infty} u(t)\theta(t)dt$ where u is a fixed real-valued function. We shall assume that the process Y can be filtered yielding a band limited version $Y_x = F^{-1}B_xF\theta + \varepsilon_x$ for any $x \in (0,x_0]$ of our choosing. The process $\varepsilon_x$ is stationary with spectral density (relative to normalized Lebesgue measure m) $f_x(\lambda) \equiv I_{[-x,x]}(\lambda)$. The mean function $\theta$ is thought to be in $\mathcal{B}_{x_0}$ but in the process of estimation we wish to guard against the possibility that $||\theta-\theta_{\mathcal{B}}||$ could be as large as $\alpha$. What is the best choice of values $\{x_1,\ldots,x_N\}$ if we employ Speckman's estimator based upon $\{Y(x_1,t),\ldots,Y(x_N,t): t \in (-\infty, +\infty)\}$?

Example 3. For each finite collection of points $\{x_1,\ldots,x_N\}$, $x_j \in (0,x_0]$ for all j, we may observe the uncorrelated stochastic processes $\{Y(x_i,t): t \in (-\infty, +\infty),\ i = 1,\ldots,N\}$ where

$$Y(x,t) = m_x(\theta)(t) + \varepsilon_x(t) \quad t \in (-\infty, +\infty),$$

$\varepsilon_x$ is a zero mean stationary process with spectral density $f_x(\lambda) = I_{[-x,x]}(\lambda)$, $\theta \in L^2$ is unknown, and $m_x = F^{-1}B_xF$. Setting $T\theta = \theta_{\mathcal{B}^\perp}$, where $\mathcal{B} = \mathcal{B}_{x_0}$, find the optimal design using Speckman's estimator with $||T\theta|| \leq \alpha$ for estimating $\tau(\theta) = \int_{-\infty}^{+\infty} u(s)\theta(s)ds$, where $u \in L^1 \cap L^2$ is real valued and continuous.

We shall prove that the optimal design places all mass at $x_0$, regardless of the function u and give the resulting best minimax linear estimator. One should note that in this problem the error process depends upon the design variable x. Our methods still apply if we replace $||m_x(\theta)||_k^2$ by $||m_x(\theta)||_{k_x}^2$ in theorem 3.2.

One can easily show that $H(K_x)$ and $L^2[-x,x]$ are isometrically isomorphic where the isometry $\psi$ satisfies

$$\psi[K_x(\cdot,t)](\lambda) = e^{-it\lambda} I_{[-x,x]}(\lambda)$$

and that $H(K_x) = F^{-1}B_xL^2 = F^{-1}B_xFL^2$. Therefore $m_x(\theta) \in H(K_x)$ for all $x \in (0,x_0]$.

Fixing N we seek $\bar{\theta}$ satisfying

$$\sup_{0<x\leq x_0} ||m_x\bar{\theta}||_{k_x}^2 + \frac{1}{N\alpha^2}||T\bar{\theta}||_2^2 =$$

(5.4)

$$\inf_{\{\theta:(u,\theta)=1\}} \sup_{0<x\leq x_0} ||m_x\theta||_{k_x}^2 + \frac{1}{N\alpha^2}||T\theta||^2.$$

We claim that

$$\bar{\theta} = \frac{u_{\mathbb{B}}+N\alpha^2 u_{\mathbb{B}}^{\perp}}{||u_{\mathbb{B}}||^2+N\alpha^2||u_{\mathbb{B}}^{\perp}||^2}. \tag{5.5}$$

In order to see this introduce

$$\Delta_1 = \{\theta: ||u_{\mathbb{B}}||\,||\theta_{\mathbb{B}}|| + ||u_{\mathbb{B}}^{\perp}||\,||\theta_{\mathbb{B}}^{\perp}|| \geq 1\}$$

Then $\Delta_0 = \{\theta: (u,\theta) = 1\} \subset \Delta_1$ so that

$$(5.6)\qquad \inf_{\Delta_1}[\sup_x ||m_x\theta||^2 + \frac{1}{N\alpha^2}||T\theta||^2] \le \inf_{\Delta_0}[\sup_x ||m_x\theta||^2 + \frac{1}{N\alpha^2}||T\theta||^2].$$

We show that $\bar{\theta} \in \Delta_1$ minimizes the left-most quantity in (5.6). However $\bar{\theta}$ is also in $\Delta_0$ showing that it also minimizes the right hand side. Since $||m_x\theta||^2_{K_x} = ||B_x F\theta||^2_2 = \int_{-x}^{x} |F\theta(t)|^2 dm(t)$ we have $\sup_{0<x\le x_0} ||m_x\theta||^2_{K_x} = ||B_{x_0} F\theta||^2_2 = ||F^{-1}B_{x_0}F\theta||^2_2$. We also have $F^{-1}B_{x_0}F\theta = \theta_{\mathbb{B}}$ since $F^{-1}B_{x_0}F$ can be shown to be the orthogonal projection onto $\mathbb{B}$. Thus the problem for $\theta \in \Delta_1$ becomes minimize $A^2 + \eta B^2$ subject to $aA + bB \ge 1$, where $a = ||u_{\mathbb{B}}||$, $b = ||u_{\mathbb{B}}^{\perp}||$, and $\eta = (N\alpha^2)^{-1}$. The solution, using calculus methods, is $A_0 = a/2\gamma$, $B_0 = b/2\eta\gamma$, and $\gamma = \frac{a^2}{2} + \frac{b^2}{2\eta}$. Choosing $\bar{\theta}$ as in (5.5) we see that indeed these norm conditions are met. Obviously $\bar{\theta} \in \Delta_0 \cap \Delta_1$ proving the claim in (5.4) that $\bar{\theta}$ is the minimizer.

Let $\xi^*$ place all mass at $x = x_0$. We now verify

$$(5.7)\qquad m^*_{x_0} m_{x_0}\bar{\theta} + \frac{1}{N\alpha^2} T^*T\bar{\theta} = (||u_{\mathbb{B}}||^2 + N\alpha^2||u_{\mathbb{B}}^{\perp}||^2)^{-1}u$$

by taking the inner product with an arbitrary $\theta$ on both sides. Working with the left hand side we have

$$(\bar{\theta}_{\mathbb{B}}, \theta_{\mathbb{B}}) + (N\alpha^2)^{-1}(\bar{\theta}_{\mathbb{B}}^{\perp}, \theta_{\mathbb{B}}^{\perp})$$

$$= \frac{(u_{\mathbb{B}}, \theta_{\mathbb{B}}) + (u_{\mathbb{B}}^{\perp}, \theta_{\mathbb{B}}^{\perp})}{||u_{\mathbb{B}}||^2 + N\alpha^2||u_{\mathbb{B}}^{\perp}||^2}$$

as required. Therefore $\xi^*$ is optimal.

By (5.7) $\tilde{\theta} = N^{-1}(||u_{\mathbb{B}}||^2 + N\alpha^2||u_{\mathbb{B}}^{\perp}||^2)\ \bar{\theta} = N^{-1}(u_{\mathbb{B}} + \alpha^2 u_{\mathbb{B}}^{\perp}$ and $m_{x_0}(\tilde{\theta}) = N^{-1}u_{\mathbb{B}}$.

The corresponding estimator is $N^{-1}\langle Y(x_0), u_{\mathcal{B}}\rangle$. If we can identify the linear operations which yield $u_{\mathcal{B}}(\cdot)$ in terms of $K_X(\cdot,t)$ then we know that the same operations on $Y(x_0)$ yield $\langle Y(x_0), u_{\mathcal{B}}\rangle$. We claim that

$u_{\mathcal{B}}(\cdot) = \int_{-\infty}^{+\infty} u(t)K_X(\cdot,t)dm(t)$. To see this consider approximating sums.

$$\int_{-\infty}^{\infty} u(t)K_X(\tau,t)dm(t) = \lim_{|\Delta t_{nj}|\to 0} \sum u(t_{nj})K_X(\tau,t_{nj})\Delta m(t_{nj})$$

$$= \lim_{|\Delta t_{nj}|\to 0} F^{-1}B_X(\sum u(t_{nj})e^{-it_{nj}} m(t_{nj}))(\tau)$$

$$= F^{-1}B_X Fu(\tau) = u_{\mathcal{B}}(\tau).$$

The minimax linear estimator of $(u,\theta)$ is therefore $\frac{1}{N}\sum_{j=1}^{N} \int_{-\infty}^{\infty} Y_j(x_0,t)u(t)dm(t)$.

This estimator has a maximum mean square error of $\frac{||u_{\mathcal{B}}||^2}{N} + \alpha^2||u_{\mathcal{B}}^{\perp}||^2$ over $\theta$ with $||\theta_{\mathcal{B}}^{\perp}|| \leq \alpha$.

If $u \in \mathcal{B}_{x_1}$ where $0 < x_1 < x_0$ then one can verify that any design whose support is contained in $[x_1,x_0]$ will be optimal.

# Bibliography

1. Fedorov, V.V. (1972). Theory of Optimal Experiments. Academic Press, New York.

2. Huber, Peter J. (1975). Robustness and designs. A Survey of Statistical Design and Linear Models. North-Holland, Amsterdam, 287-303.

3. Kiefer, J. (1980). Designs for extrapolation when bias is present. Multivariate Analysis V. North-Holland, Amsterdam, 79-93.

4. Li, Ker-Chau (1981). Robust regression designs when the design space consists of finitely many points. Purdue Mimeo Series 81-45.

5. Li, K.C. and Notz, W. (1980). Robust designs for nearly linear regression. Purdue Mimeo Series 80-3.

6. Marcus, M. and Sacks, J. (1976). Robust designs for regression problems. Statistical Decision Theory and Related Topics II, Academic Press, New York, 245-268.

7. Notz, W. (1980). Optimal robust designs for some regression problems. Purdue Mimeo Series 80-29.

8. Pesotchinsky, L. (1979). Robust designs and optimality of least squares for regression problems. Manuscript.

9. Slepian, D. and Pollak, H.O. (1961). Prolate spheroidal wave functions Fourier analysis and uncertainty - I. Bell System Technical Journal.

10. Speckman, P. (198 ). Minimax estimates of linear functionals in a Hilbert space. To appear in Ann. Statist.

11. Spruill, Carl (1980). Optimal designs for second order processes with general linear means. Ann. Statist. 8 652-663.

12. Spruill, M.C. (1981). Traversing the smallest possible corridor under velocity and energy constraints. Purdue Mimeo Series 81-52.

13. Spruill, Carl and Studden, W.J. (1978). Optimum designs when the observations are second order processes. J. Multivariate Analysis.