Optimal designs for estimating the slope of a polynomial regression

by

V. N. Murty and W. J. Studden*

Penn State University, Capitol Campus

and Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 271

November, 1971

---

* This research was partially supported by the National Science Foundation, Grant No. GP 20306.

Abstract - Optimal designs for estimating the slope of a polynomial regression - By V. N. Murty and W. J. Studden

This paper is divided into two parts. Part one consists of a brief review of the general design problem, emphasizing the Kiefer and Wolfowitz (1959) characterization of c-optimal designs and the Federov (1971) characterization of L-optimal designs. In the second part we present the designs for estimating the slope of a second and third degree polynomial at a fixed point of the experimental region with minimum variance. Designs are also considered for minimizing the integrated variance of the estimated slope for second and third degree polynomials, where the integration is carried out with respect to a fixed probability measure over the experimental region or over an extended domain. Finally, for the second degree polynomial we present the design that minimizes the integrated variance of the estimated regression function.

Optimal designs for estimating the slope of a polynomial regression

by

V. N. Murty and W. J. Studden
Penn State University, Capitol Campus
and Purdue University

§1. Introduction. The design problem under discussion is as follows. Let $f' = (f_0, f_1, \ldots, f_n)$ denote an $(n+1)$-vector of continuous functions defined on a compact set X. The points of X are referred to as the possible levels of feasible experiments and the variable $x \in X$ is sometimes called the control variable. For each level $x \in X$ some experiment may be performed whose outcome $Y(x)$ is a random observation with mean value

$$(1.1) \qquad E[Y(x)] = \sum_{i=0}^{n} \theta_i f_i(x)$$

and variance σ^2 independent of x. The functions $f_0, f_1, \ldots, f_n$ are called the regression functions and are known to the experimenter. The regression coefficients or parameters $\theta_0, \theta_1, \ldots, \theta_n$ and σ^2 are unknown. On the basis of N uncorrelated observations we wish to estimate some function of the parameters $\theta_0, \theta_1, \ldots, \theta_n$.

An experimental design specifies a probability measure μ (usually discrete) on X. The associated experiment involves taking observations at the level x proportional to μ. Thus if μ assigns mass $p_0, p_1, \ldots, p_r$ to $x_0, x_1, \ldots, x_r$ and $NP_i = n_i$ are integers the experimenter takes n_i observations at x_i. Designs with Np_i not equal to an integer can in practice only be approximated.

If the unknown parameter vector $\theta' = (\theta_0, \theta_1, \ldots, \theta_n)$ is estimated by least squares then the covariance matrix of the estimates $\hat{\theta}$ is given by

$$(1.2) \qquad E(\hat{\theta}-\theta)(\hat{\theta}-\theta)' = \frac{\sigma^2}{N} \cdot M^{-1}(\mu)$$

The matrix $M(\mu) = \int_X f(x)\, f'(x)\, d\mu(x)$ is called the <u>information matrix</u> of the design μ.

The variance of the least square estimator of the regression function at the point $x \in X$ is proportional to

$$(1.3) \qquad f'(x)\, M^{-1}(\mu)\, f(x) = \text{tr}\, M^{-1}(\mu)\, f(x)\, f'(x)$$

where tr denotes the trace of a matrix. The variance of the least square estimator of a linear form $(c,\theta) = \sum_{i=0}^{n} c_i \theta_i$ is proportional to

$$(1.4) \qquad c'\, M^{-1}(\mu)\, c = \text{tr}\, M^{-1}(\mu)\, c\, c' \quad .$$

Let $g' = (g_0, g_1, \ldots, g_n)$ denote the $(n+1)$-vector, where $g_i = \frac{d}{dx} f_i$. The slope of the regression function (1.1) at a point $x \in X$ is given by

$$(1.5) \qquad \sum_{i=0}^{n} \theta_i g_i = (\theta, g) \quad .$$

The variance of the least square estimator of (1.5) using the design μ is proportional to

$$(1.6) \qquad g'(x)\, M^{-1}(\mu)\, g(x) = \text{tr}\, M^{-1}(\mu)\, g(x)\, g'(x) \quad .$$

The integrated variance with respect to a fixed probability measure σ, of the least square estimator of the regression function (1.1) is proportional to (using design μ)

$$(1.7) \qquad \int_X f'(x)\, M^{-1}(\mu)\, f(x)\, d\sigma(x)$$

$$= \operatorname{tr} M^{-1}(\mu)\, M(\sigma)$$

where $$M(\sigma) = \int_X f(x)\, f'(x)\, d\sigma(x) \quad .$$

The integrated variance of the least square estimator of the estimated slope of the regression function is proportional to (using design μ)

$$(1.8) \qquad \int_X g'(x)\, M^{-1}(\mu)\, g(x)\, d\sigma(x)$$

$$= \operatorname{tr} M^{-1}(\mu)\, C$$

where $$C = \int_X g(x)\, g'(x)\, d\sigma(x) \quad .$$

§2. <u>C-optimal and L-optimal designs</u>. A design μ_1^* is said to be a c-optimal design for estimating the linear form (c,θ) if it minimizes (1.4) i.e.

$$c'\, M^{-1}(\mu_1^*)c = \min_{\mu} c'\, M^{-1}(\mu)\, c \quad .$$

A design μ_2^* that minimizes (1.7) is called an L-optimal design. Studden (1971) denotes this by I_σ-optimal design. Designs minimizing (1.8) also come under the general category of L-optimal designs of Federov (1971).

For any vector $c \neq (0,0,\ldots,0)$ we define the determinants $D_\nu(c)$, where

$$(2.1) \qquad D_\nu(c) = \begin{vmatrix} f_0(s_0) & \cdots & f_0(s_{\nu-1}) & f_0(s_{\nu+1}) & \cdots & f_0(s_n) & c_0 \\ f_1(s_0) & \cdots & f_1(s_{\nu-1}) & f_1(s_{\nu+1}) & \cdots & f_1(s_n) & c_1 \\ \vdots & & & & & \vdots & \vdots \\ f_n(s_0) & \cdots & f_n(s_{\nu-1}) & f_n(s_{\nu+1}) & \cdots & f_n(s_n) & c_n \end{vmatrix}$$

$$\nu = 0, 1, \ldots, n$$

and $s_0, s_1, \ldots, s_n$ are $(n+1)$ points of the experimental region on which a design μ concentrates mass, and $s_0 < s_1 < s_2 \ldots < s_n$. The sign of $D_\nu(c)$ will be denoted by $d_\nu(c)$; if $D_\nu(c) = 0$ the sign may be defined as -1 or $+1$. We denote by $\ell_\nu(x)$ the Lagrange basis functions associated with the set of $(n+1)$ points $s_0, s_1, \ldots, s_n$ for the original regression functions $f_0, f_1, \ldots, f_n$ and are given by

$$(2.2) \qquad \ell_\nu(x) = \frac{\begin{vmatrix} f_0(s_0) & \cdots & f_0(s_{\nu-1}) & f_0(s_{\nu+1}) & \cdots & f_0(s_n) & f_0(x) \\ f_1(s_0) & \cdots & f_1(s_{\nu-1}) & f_1(s_{\nu+1}) & \cdots & f_1(s_n) & f_1(x) \\ \vdots & & & & & \vdots & \vdots \\ f_n(s_0) & \cdots & f_n(s_{\nu-1}) & f_n(s_{\nu+1}) & \cdots & f_n(s_n) & f_n(x) \end{vmatrix}}{\begin{vmatrix} f_0(s_0) & \cdots & f_0(s_{\nu-1}) & f_0(s_\nu) & \cdots & f_0(s_n) \\ f_1(s_0) & \cdots & f_1(s_{\nu-1}) & f_1(s_\nu) & \cdots & f_1(s_n) \\ \vdots & & \vdots & \vdots & & \vdots \\ f_n(s_0) & \cdots & f_n(s_{\nu-1}) & f_n(s_\nu) & & f_n(s_n) \end{vmatrix}}$$

Note that $\ell_\nu(s_j) = \begin{cases} 1 & \text{if } \nu = j \\ 0 & \text{if } \nu \neq j \end{cases}$

$$\nu = 0,1,\ldots, n; \quad \text{and} \quad j = 0,1,\ldots, n \quad .$$

We make use of the following theorem due to Karlin and Studden (1966) [see also Studden 1968] which is very closely related to the results of Kiefer and Wolfowitz (1965) in obtaining c-optimal designs.

Let R denote the class of vectors c such that $\varepsilon D_\nu(c) \geq 0$ for $\nu = 0,1,\ldots, n$ where ε is fixed at $+1$ or -1 for a given vector c and let S denote the class of vectors c for which $\varepsilon(-1)^\nu D_\nu(c) \leq 0$ for $\nu = 0,1,\ldots,n$. The theorem referred to above says:

Theorem 2.1

If $\{f_i\}_0^n$ is a Tchebycheff system on X and there exists a linear combination of the f_i's, such that it is $\equiv 1$ (linear combination is denoted by $U(x)$) then

(a) For any design μ, the variance of the least square estimator of the linear form (c,θ) using design μ is always greater than or equal to $[W(c)]^2$ for $c \in R$ and $[U(c)]^2$ for $c \in S$, where

$$W(c) = \sum_{i=0}^{n} a_i c_i; \qquad U(c) = \sum_{i=0}^{n} b_i c_i$$

the a_i's being the coefficients in the unique linear combination of f_i's that oscillates between -1 and $+1$ attaining these extreme values with alternating signs at $(n+1)$ points $s_0, s_1, \ldots, s_n$ called the Tchebycheff points. The b_i's are the coefficients in the linear combination of f_i's that is $\equiv 1$.

(b) The variance of the least square estimator of (c,θ) will be equal to $[W(c)]^2$ for $c \in R$ and $[U(c)]^2$ for $c \in S$ if the design $\mu = \mu_1^*$ concentrates mass at the points s_ν; $\nu = 0,1,\ldots, n$ with weights

$$p_\nu = |D_\nu(c)| \Big/ \sum_{\nu=0}^{n} |D_\nu(c)|$$

(c) The design μ_1^* is the only design supported on $s_0 < s_1 \ldots < s_n$ for which the variances are equal to $[W(c)]^2$ or $[U(c)]^2$. If $c \in R$, μ_1^* is unique.

For obtaining the L-optimal designs we mainly use the following lemma due to Federov (1971) if we are interested in minimizing (1.7). Before stating this lemma we first note that the expression $\operatorname{tr} M^{-1}(\mu)\, M(\sigma)$ is invariant under basis change of the regression functions, i.e. if instead of the regression functions $f_0, f_1, \ldots, f_n$ which we assumed to be linearly independent we take as our regression functions another set of $(n+1)$ linearly independent functions, which are linear combinations of these, and compute this trace we get the same number. Thus if a design μ^* concentrates its mass on $(n+1)$ points $s_0 < s_1, \ldots < s_n$ and we consider as our regression functions $\ell_0(x)$, $\ell_1(x), \ldots, \ell_n(x)$ given at (2.2) and call

$$M_\ell(\mu^*) = \int_X \ell(x)\ \ell'(x)\ d\mu^*(x)$$

$$M_\ell(\sigma) = \int_X \ell(x)\ \ell'(x)\ d\sigma(x)$$

then

$$\operatorname{tr} M^{-1}(\mu^*)\ M(\sigma) = \operatorname{tr} M_\ell^{-1}\ (\mu^*)\ M_\ell(\sigma) \quad .$$

Moreover $M_\ell(\mu)$ is a diagonal matrix with elements $p_0, p_1, \ldots, p_n$ the weights of μ^* on its diagonal so that

$$\operatorname{tr} M^{-1}(\mu^*)\ M(\sigma) = \sum_{i=0}^{n} k_i^2/p_i$$

where

$$k_i = \int_X \ell_i^2(x)\ d\sigma(x)$$

Federov's Lemma: If the design μ_2^* that minimizes (1.7), for a given σ, concentrates mass on $s_0, s_1, \ldots, s_n$ then the corresponding weights are proportional to $\sqrt{k_\nu}$; $\nu = 0, 1, \ldots, n$.

To obtain designs that minimize (1.8) we first note that the matrix

$$C = \int_X f(x)\ g'(x)\ d\sigma(x)$$

is a positive semidefinite symmetric matrix and hence can be written as $C = A\ A'$

where A is an $(n+1)$ square matrix. So the problem reduces to minimizing $\text{tr}\, M^{-1}(\mu)\, A\, A'$ and hence we can use the following theorem of Studden (1970) that characterizes such designs.

Theorem 2.2

A design μ_3^* concentrating mass at $(n+1)$ points $s_0, s_1, \ldots, s_n$ minimizes (1.8) if and only if there exists a matrix B such that

(i) $\ell'(x)\, B_0\, B_0'\, \ell(x) \leq 1$ for all x

(ii) $A = F\, B$

where $\ell'(x)$ is the row vector of Lagrange basis functions, F is the matrix with columns $f'(s_\nu)$ and is assumed to be non-singular, and the matrix B_0 is obtained from B by taking each non zero row of B and normalizing it so as to make its length unity, i.e. if $b_{i0}, b_{i1}, \ldots, b_{in}$ is the ith row of B, then the ith row of B_0 is given by $b_{i0}/|b_i|, b_{i1}/|b_i|, \ldots, b_{in}/|b_i|$ where $|b_i| = \sqrt{b_{i0}^2 + b_{i1}^2 + \ldots + b_{in}^2}$. The weights of μ_3^* are proportional to the lengths of the rows of B.

The following theorem of Elfving will also be needed.

Elfving Theorem

The design μ minimizes $c'\, M^{-1}(\mu)\, c$ if and only if there exists $\varepsilon_\nu = \pm 1$ such that $\beta c = \sum \varepsilon_\nu p_\nu f(x_\nu)$ and βc is in the boundary of R. Here μ concentrates mass p_ν at x_ν $\nu = 1, 2, \ldots$ and R is the convex hull of the set $\{\pm f(x) | x \in X\}$.

§3. Optimal designs for estimating the slope of a second and third degree polynomial regression. From now on we take $X = [-1,1]$ and our vector of regression functions $f'(x)$ is either

(3.1) $$f'(x) = (1, x, x^2)$$

or

(3.2) $$f'(x) = (1, x, x^2, x^3)$$

so that

(3.3) $$g'(x) = (0, 1, 2x)$$

or

(3.4) $$g'(x) = (0, 1, 2x, 3x^2) \quad .$$

3.1 Quadratic regression

To obtain the design μ_1^* that minimizes (1.4) we now take $(c_0,c_1,c_2) = (0, 1, 2x)$, x a fixed point in $[-1,1]$, we use the Theorem 2.1, after noting that $(1, x, x^2)$ is a Tchebycheff system on $[-1,1]$ and the unique linear combination of $1, x, x^2$ that oscillates between -1, and 1 is $T_2(x) = 2x^2 - 1$ which attains its maximum with alternating signs at $s_0 = -1$, $s_1 = 0$, and $s_2 = +1$. So we compute $D_0(c)$, $D_1(c)$, and $D_2(c)$ which are given by

(3.5) $$D_0(c) = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2x \end{vmatrix} = 2x - 1$$

$$D_1(c) = \begin{vmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \\ 1 & 1 & 2x \end{vmatrix} = 4x$$

$$D_2(c) = \begin{vmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 2x \end{vmatrix} = 2x + 1$$

so that $c \in R$ (with $\varepsilon = 1$) if $2x \geq 1$ or $2x \leq -1$ (with $\varepsilon = -1$).

Therefore the design μ_1^* concentrating mass at the Tchebycheff points i.e. $\{x \mid |T_2(x)| = 1\}$ is the unique design for estimating the slope of a quadratic regression function with minimum variance with weights

$$(3.6) \qquad p_\nu = |D_\nu(c)| \Big/ \sum_{\nu=0}^{2} |D(c)| \; ; \qquad \nu = 0,1,2$$

at a point x where $x \geq 1/2$ or $x \leq -1/2$. From (3.5) these weights are

$$p_0 = \frac{1}{4} - \frac{1}{8x} , \quad p_1 = \frac{1}{2} , \quad p_3 = \frac{1}{4} + \frac{1}{8x} \; ;$$

If $-1/2 < x < 1/2$, we cannot apply Theorem 2.1. It can be checked that, if $-1/2 < x < 0$, the design μ_1^* concentrates mass at $s_0 = -1$ and $2x + 1 = s_1$ with equal weights. A direct appeal to Elfving's Theorem [see Karlin and Studden (1966)] will prove this assertion. Similarly, if $0 < x < 1/2$, μ_1^* concentrates mass at $s_0 = 2x-1$ and $s_1 = 1$ with equal weights. The designs obtained above are summarized in the following table.

Table 3.1

Optimal designs for estimating slope

with a quadratic regression

Serial No.	Points at which slope is estimated	Optimal design concentrates mass at	Optimal weights
1	$x \in [-1, -1/2]$	$s_0 = 1;\ s_1 = 0;\ s_2 = 1$	see (3.6)
2	$x \in (-1/2, 0)$	$s_0 = -1;\ s_1 = 2x + 1$	$p_0 = p_1 = 1/2$
3	$x \in [0, 1/2)$	$s_0 = 2x - 1;\ s_1 = 1$	$p_0 = p_1 = 1/2$
4	$x \in [\frac{1}{2}, 1]$	$s_0 = -1;\ s_1 = 0;\ s_2 = 1$	see (3.6)

To obtain the design that minimizes (1.8) when the regression is quadratic we first note that

$$\int_X g'(x)\, M^{-1}(\mu)\, g(x)\, d\sigma(x) = \text{tr } M^{-1}(\mu)\, C$$

where

$$C = \int_X g(x)\, g'(x)\, d\sigma(x)$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2\mu_1 \\ 0 & 2\mu_1 & 4\mu_2 \end{bmatrix}$$

and

$$\text{tr } M^{-1}(\mu)\, C = \text{tr } M_\ell^{-1}(\mu)\, T\, C\, T'$$

where

$$T = \begin{bmatrix} \frac{s}{2(1+s)} & -1/2 & \frac{1}{2(1+s)} \\ \frac{1}{1-s^2} & 0 & -\frac{1}{1-s^2} \\ -\frac{s}{2(1-s)} & 1/2 & \frac{1}{2(1-s)} \end{bmatrix}$$

$$M_\ell^{-1}(\mu) = \begin{bmatrix} 1/p_0 & 0 & 0 \\ 0 & 1/p_1 & 0 \\ 0 & 0 & 1/p_2 \end{bmatrix}$$

where p_0, p_1, p_2 are the weights that the design μ concentrates at -1, s, and 1 respectively. Thus if we denote the matrix $T\, C\, T' = K$

$$\text{tr } M^{-1}(\mu)\, C = \sum_{i=0}^{2} k_i/p_i$$

which is minimized when p_i is proportional to $\sqrt{k_i}$. In our case

$$k_1 = \frac{1}{4} - \frac{\mu_1}{1+s} + \frac{\mu_2}{(1+s)^2}$$

$$k_2 = \frac{4\mu_2}{(1-s^2)^2}$$

$$k_3 = \frac{1}{4} + \frac{\mu_1}{1-s} + \frac{\mu_2}{(1-s)^2}$$

where $\mu_i = \int x^i \, d\sigma(x)$, $i = 1,2$. Thus the minimum value of $\mathrm{tr}\, M^{-1}(\mu)\, C$ is given by

(3.7) $$\mathrm{tr}\, M^{-1}(\mu)\, C = \left(\sum_{i=0}^{2} \sqrt{k_i} \right)^2$$

and if the value s is to minimize (3.7) then s must be a solution of

(3.8) $$\frac{k_1'}{\sqrt{k_1}} + \frac{k_2'}{\sqrt{k_2}} + \frac{k_3'}{\sqrt{k_3}} = 0$$

where $k_i' = dk_i/ds$. This equation reduces to

$$(1-s)^2[\mu_1 - 2\mu_2 + s\mu_1]\, g^{-1/2}(s) + 4s\sqrt{\mu_2}$$

$$+ (1+s)^2[\mu_1 + 2\mu_2 - s\mu_1] h^{-1/2}(s) = 0$$

where $$g(s) = s^2 + 2s(1-2\mu_1) + (1-4\mu_1+4\mu_2)$$

and $$h(s) = s^2 - 2s(1+2\mu_1) + (1+4\mu_1+4\mu_2) .$$

There does not appear to be closed expression for s. The value $s = 0$ is a solution of (3.8) if $\mu_1 = 0$. Note that $s = 0$ is also a solution if $\mu_1 = x$, $\mu_2 = x^2$, (σ concentrates all mass at x) and $|\mathbf{x}| \geq 1/2$.

The result for $\mu_1 = 0$ can also be analysed using Theorem 2.2. In this case the matrix C becomes a diagonal matrix and could be written as $\Lambda \Lambda'$

where

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2\sqrt{\mu_2} \end{bmatrix}$$

Using Theorem (2.2) with $s_0 = -1$, $s_1 = 0$, $s_2 = 1$

$$\ell'(x) = \left(\frac{x^2-x}{2}, \quad 1-x^2, \quad \frac{x^2+x}{2} \right)$$

$$B_0 = \begin{bmatrix} 0 & -1/\sqrt{1+4\mu_2} & 2\sqrt{\mu_2}/\sqrt{1+4\mu_2} \\ 0 & 0 & -1 \\ 0 & 1/\sqrt{1+4\mu_2} & 2\sqrt{\mu_2}/\sqrt{1+4\mu_2} \end{bmatrix}$$

and

$$F = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

one can easily check that all the conditions are satisfied and hence the optimal design that minimizes the integrated variance of the estimated slope of a quadratic regression with respect to an arbitrary but fixed measure σ that satisfies $\int_X x \, d\sigma(x) = 0$ concentrates mass at $s_0 = -1$, $s_1 = 0$, and $s_2 = 1$ with weights proportional to $\frac{1}{2}(1+4\mu_2)^{1/2}$, $2\sqrt{\mu_2}$ and $\frac{1}{2}(1+4\mu_2)^{1/2}$ respectively. If σ is the uniform measure i.e. $d\sigma(x) = dx$ then the weights are 1/4, 1/2, 1/4 which was also observed by Ott and Mendenhall (1970).

3.2 Cubic Regression:

To obtain the design μ_1^* that minimizes (1.4) we now take

$$(c_0, c_1, c_2) = (0, 1, 2x, 3x^2)$$

where x is a fixed point in $[-1,1]$, and use Theorem 2.1, noting that $(1, x, x^2, x^3)$ is a Tchebycheff system on $[-1,1]$ and the unique linear combination of $1, x, x^2, x^3$ that oscillates between -1, and $+1$ is

$$T_3(x) = 4x^3 - 3x$$

which attains its maximum with alternating signs at $s_0 = -1$, $s_1 = -1/2$; $s_2 = 1/2$, and $s_3 = 1$. Computation yields

$$(3.9)\quad D_0(c) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -1/2 & 1/2 & 1 & 1 \\ 1/4 & 1/4 & 1 & 2x \\ -1/8 & 1/8 & 1 & 3x^2 \end{bmatrix} = \frac{1}{16}[36x^2-24x-3]$$

$$D_1(c) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & 1/2 & 1 & 1 \\ 1 & 1/4 & 1 & 2x \\ -1 & 1/8 & 1 & 3x^2 \end{bmatrix} = \frac{1}{2}[9x^2-3x-3]$$

$$D_2(c) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & -1/2 & 1 & 1 \\ 1 & 1/4 & 1 & 2x \\ -1 & -1/8 & 1 & 3x^2 \end{bmatrix} = \frac{1}{2}[9x^2+3x-3]$$

$$D_3(c) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & -1/2 & 1/2 & 1 \\ 1 & 1/4 & 1/4 & 2x \\ -1 & -1/8 & 1/8 & 3x^2 \end{bmatrix} = \frac{1}{36}[36x^2+24x-3]$$

so that $c \varepsilon R$ if (a) $x \varepsilon [-1, \frac{-2-\sqrt{7}}{6}]$ or (b) $x \varepsilon [\frac{2-\sqrt{7}}{6}, \frac{-2+\sqrt{7}}{6}]$ or (c) $x \varepsilon [\frac{2+\sqrt{7}}{6}, 1]$. Hence the design μ_1^* concentrating mass at the Tchebycheff points i.e. $\{x \mid |T_3(x)| = 1\}$ is the unique design for estimating the slope of a cubic regression function with minimum variance with weights

$$(3.10) \qquad p_\nu = D_\nu(c) \Big/ \sum_{\nu=0}^{3} |D_\nu(c)| ; \quad \nu = 0,1,2,3$$

at s_γ where $s_0 = -1$, $s_1 = -1/2$, $s_2 = 1/2$ and $s_3 = +1$ if the fixed point x is in (a), (b) or (c). The optimal design for the other cases are given in **Table** 3.2. These results can be verified using Elfving's Theorem. They were obtained by considering the cubic polynomial lying between ± 1 on $[-1, 1]$ with a maximum derivative at the point x. The x values where the resulting polynomial touches ± 1 support the optimal design.

To minimize (1.8), when the regression is cubic we present below some computer results obtained when the measure σ is of the following type.

$$d\sigma(x) = k \cdot (1+x)^{\alpha-1} (1-x)^{\alpha-1} dx \quad .$$

The corresponding optimal design is on $s_0 = -1$, $s_1 = -z$, $s_2 = +z$, $s_3 = +1$ with weights q, p, p, q respectively so that

$$2(p+q) = 1 \quad .$$

The results obtained for this case are presented in Table 3.3.

Table 3.2

Optimal designs for estimating slope with a cubic regression

Serial No.	Point at which slope is estimated	Optimal design concentrates mass at	Optimal weights
1	$x\varepsilon[-1,\frac{-2-\sqrt{7}}{6}]$	$s_0=-1, s_1=-1/2, s_2=1/2, s_3=1$	p_0, p_1, p_2, p_3 [see (3.10)]
2	$x\varepsilon(\frac{-2-\sqrt{7}}{6},\frac{-1-2\sqrt{7}}{9})$	$s_0=-1, s_1=\frac{y-2}{3}; s_2=y$ $y=x(4+\sqrt{7})+(3+\sqrt{7})$	$\frac{2\sqrt{7}+8}{27}, 1/2, \frac{11-4\sqrt{7}}{54}$
3	$x\varepsilon[\frac{-1-2\sqrt{7}}{9}, \frac{1-2\sqrt{7}}{9}]$	$s_0=-1, s_1=y; s_2=1$ $y=3x+\frac{2}{3}\sqrt{7}$	$p_0=(1+y-2x)(y-1)/8x$ $p_1=1/2; p_2=(2x-y+1)(1+y)/8x$
4	$x\varepsilon(\frac{1-2\sqrt{7}}{9}, \frac{2-\sqrt{7}}{6})$	$s_0=y, s_1=\frac{y+2}{3}, s_2=1$ $y=x(4+\sqrt{7})-(3+\sqrt{7})$	$\frac{11-4\sqrt{7}}{54}, 1/2, \frac{2\sqrt{7}+8}{54}$
5	$x\varepsilon[\frac{2-\sqrt{7}}{6}\quad\frac{-2+\sqrt{7}}{6}]$	$s_0=-1, s_1=-1/2, s_1=1/2, s_3=1$	p_0, p_1, p_2, p_3; [see (3.10)]
6	$x\varepsilon(\frac{-2+\sqrt{7}}{6}, \frac{-1+2\sqrt{7}}{9})$	$s_0=y, s_1=\frac{y+2}{3}, s_2=1$ $y=x(4+\sqrt{7})-(3+\sqrt{7})$	$\frac{11-4\sqrt{7}}{54}, 1/2, \frac{2\sqrt{7}+8}{54}$
7	$x\varepsilon[\frac{-1+2\sqrt{7}}{9}, \frac{1+2\sqrt{7}}{9}]$	$s_0=-1, s_1=y, s_2=1$ $y=3x-\frac{2}{3}\sqrt{7}$	same weights as in 3.
8	$x\varepsilon(\frac{1+2\sqrt{7}}{9}, \frac{2+\sqrt{7}}{6})$	$s_0=-1, s_1=\frac{y-2}{3}, s_2=y$ $y=x(4+\sqrt{7})+(3+\sqrt{7})$	same weights as in 2.
9	$x\varepsilon[\frac{2+\sqrt{7}}{6}, 1]$	$s_0=-1, s_1=-1/2, s_2=1/2,$ $s_3=1$	same weights as in 1.

Table 3.3

Optimal design that minimizes

the integrated variance of the estimated slope

with a cubic regression, when $d\sigma(x) = k(1+x)^{\alpha-1}(1-x)^{\alpha H}\,dx$.

α	z	p	q
0.1	0.453	0.291	0.209
0.5	0.445	0.288	0.212
1.0	0.442	0.291	0.209
1.5	0.447	0.299	0.201
2.0	0.456	0.309	0.191
2.5	0.464	0.319	0.181
3.0	0.472	0.328	0.172
4.0	0.484	0.345	0.155
5.0	0.492	0.358	0.142

4. In this section we present some simple results concerning the design that minimizes (1.7) when the regression is quadratic. Here we need search for a design μ that concentrates mass at -1, s, and 1 with weights p_0, p_1, and p_2 respectively. We can now use the Federov Lemma and obtain the point s, as the root of the equation

$$\frac{k_1'}{\sqrt{k_1}} + \frac{k_2'}{\sqrt{k_2}} + \frac{k_3'}{\sqrt{k_3}} = 0 \tag{4.1}$$

where

$$k_i' = \frac{dk_i}{ds} \qquad i = 1,2,3$$

and

$$k_i = \int_X \ell_i^2(x)\, d\sigma(x)$$

$$\begin{aligned} \ell_1(x) &= (x-s)(x-1)/2(1+s) \\ \ell_2(x) &= (1-x^2)/(1-s^2) \\ \ell_3(x) &= (x+1)(x-s)/2(1-s) \end{aligned} \tag{4.2}$$

$$\begin{aligned} \frac{k_1'}{\sqrt{k_1}} &= \frac{1}{(1+s)^2} \frac{[-\mu_4+\mu_3(1+s)+\mu_2(1-s)-\mu_1(1+s)+s]}{[\mu_4-2\mu_3(1+s)+\mu_2(1-4s+s^2)-\mu_1\cdot 2s(1+s)+s^2]^{1/2}} \\ \frac{k_2'}{\sqrt{k_2}} &= \frac{4s}{(1-s^2)^2} \cdot \frac{1}{(1-2\mu_2+\mu_4)^{1/2}} \\ \frac{k_3'}{k_3} &= \frac{1}{(1-s)^2} \frac{[\mu_4+\mu_3(1-s)-\mu_2(1+s)-\mu_1(1-s)+s]}{[\mu_4+2(1-s)\mu_3+\mu_2(1-4s+s^2)-2s(1-s)\mu_1+s^2]^{1/2}} \end{aligned} \tag{4.3}$$

$s = 0$ is a root of (4.1) if

$$\frac{(\mu_4-\mu_3-\mu_2+\mu_1)^2}{\mu_4-2\mu_3+\mu_2} = \frac{(\mu_4+\mu_3-\mu_2-\mu_1)^2}{\mu_4+2\mu_3+\mu_2}$$

which is true if $\mu_1 = \mu_3 = 0$. In other cases it is not easy to give a closed form expression for s.

References

Elfving, G. (1952). Optimum allocation in linear regression theory. Ann. Math. Statist. 23, 255-262.

Federov, V. V. (1971). Theory of Optimal Experiments, (to appear), Academic Press, New York.

Karlin, S. and Studden, W. J. (1966). Optimal experimental designs. Ann. Math. Statist. 37, 783-815.

Kiefer and Wolfowitz (1959). Optimum designs in regression problems. Ann. Math. Statist. 30, 271-294.

Kiefer and Wolfowitz (1965). On a theorem of Hoel and Levine on extrapolation. Ann. Math. Statist. 36, 1627-1655.

Ott and Mendenhall (1970). Designs for estimating the slope of a second order linear model. Tech. Report No. 13, Dept. of Stat. University of Florida, Gainesville. (To appear in Technometrics).

Studden, W. J. (1968). Optimal designs on Tchebycheff points. Ann. Math. Statist. 39, 1435-1447.

Studden, W. J. (1971). Elfving's theorem and optimal designs for quadratic loss, Ann. Math. Statist. 42, 1613-1621.

Studden, W. J. (1971). Optimal designs and spline regression. Dept. of Stat., Purdue University, Mimeo Series #257.