The Annals of Statistics
1979, Vol. 7, No. 1, 77–95

ROBUSTNESS OF DESIGN AGAINST AUTOCORRELATION IN TIME I: ASYMPTOTIC THEORY, OPTIMALITY FOR LOCATION AND LINEAR REGRESSION

BY P. J. BICKEL[1] AND AGNES M. HERZBERG

University of California, Berkeley and
Imperial College, London

A new asymptotic theory for studying the effect of dependence of the observations in experimental design for the linear model is developed. The uniform design is shown to be asymptotically optimal in a strong sense for estimating location and in a weaker sense for estimating the slope of a straight line regression. Numerical results supporting the asymptotics appear in a companion paper.

1. Introduction. Suppose observations can be taken on a variable y at N time points, $-T \leqslant t_1 \leqslant \cdots \leqslant t_N \leqslant T$. Suppose also that an observation at time t can be written

$$Y(t) = \beta_1 f_1(t) + \cdots + \beta_p f_p(t) + \varepsilon(t), \tag{1.1}$$

where the $f_j(t)$ are known functions, the β_j are unknown parameters ($j = 1, \ldots, p$) and $\varepsilon(t)$ is a random error with 'centre' 0. Much attention has been paid to the problem of selecting the t_i, i.e., designing the experiment, in order to minimize some measure of the variability of the estimates of the β_i. If the errors corresponding to different observations are assumed to be independent and identically distributed as $N(0, \sigma^2)$, the problem becomes a special case of the general theory of optimal design developed by Kiefer and others; see, for example, Kiefer (1974). It has, however, been known for some time that the solutions obtained in this way are sensitive to departures from the assumed linear form of the mean response (Huber, 1975) and the assumed independence of the errors. Here the latter issue will be pursued, namely robustness against dependence of the errors.

Suppose that the observations $Y_i(t_i)$ or the corresponding errors $\varepsilon_i(t_i)$ ($i = 1, \ldots, N$) have a joint normal distribution with

$$E\{\varepsilon_i(t_i)\} = 0, \tag{1.2}$$

$$\operatorname{Var}\{\varepsilon_i(t_i)\} = \sigma^2, \tag{1.3}$$

$$\operatorname{corr}\{Y_i(t_i), Y_j(t_j)\} = \gamma\rho(t_i - t_j) \qquad i \neq j, \tag{1.4}$$

where $0 \leqslant \gamma \leqslant 1$ and $\rho(\cdot)$ is the correlation function of a nondegenerate stationary

---

[1]This research was prepared with the support of National Science Foundation Grant No. MCS76-10238 and U.S. Office of Naval Research Contract No. N00014-75-C-0444/NR042-036. A portion of this work was done while the author held a NATO Senior Science Fellowship at Imperial College, London.

AMS 1970 *subject classifications.* 6262, 6215, 6285.

Key words and phrases. Experimental design, autocorrelation in time, optimal, asymptotic.

process, i.e., $\rho(\cdot)$ is a positive definite function which is symmetric and $\rho(0) = 1$. This model can be motivated as follows.

If $\varepsilon_1(s)$ and $\varepsilon_2(t)$ are the errors of two observations taken at times s and t, then

$$\text{(1.5)} \qquad \begin{aligned} \varepsilon_1(s) &= \varepsilon'(s) + \varepsilon_1'', \\ \varepsilon_2(t) &= \varepsilon'(t) + \varepsilon_2'', \end{aligned}$$

where $\varepsilon'(\cdot)$ is a stationary Gaussian process with mean 0, correlation function ρ and variance $\gamma\sigma^2$ while the ε_i'' are independent $N\{0, (1 - \gamma)\sigma^2\}$ variables. Note that even if $s = t$ our observations need not be identical. Such a model naturally suggests itself in various situations.

An important class of examples pointed out by Morrison (1970) includes repeated measurements of a biological variable on single individuals. Another important class includes the situation in which the same observer makes repeated measurements. The evidence for dependence in such cases is very strong; see, for example, Pearson's data as discussed in Jeffreys (1939). Other related work is that of Hoel (1958), Bloomfield and Watson (1975) and Knott (1975).

If the form of (1.4) is assumed known, the design problem becomes one of minimizing some measure of the size of the variance covariance matrix $\mathbf{\Sigma}$ of the best linear unbiased estimators of $\beta_1, \ldots, \beta_p$, where $\mathbf{\Sigma}$ is given by

$$\text{(1.6)} \qquad \mathbf{\Sigma} = (\mathbf{F}^T\mathbf{U}^{-1}\mathbf{F})^{-1}.$$

Further

$$\text{(1.7)} \qquad \mathbf{F}^T = \{f_i(t_j)\} \qquad i = 1, \ldots, p, j = 1, \ldots, N$$

is a $p \times N$ matrix and

$$\text{(1.8)} \qquad \mathbf{U} = \{\gamma\rho(t_i - t_j) + (1 - \gamma)\delta_{ij}\} \qquad i, j = 1, \ldots, N$$

is an $N \times N$ matrix.

The dependence of $\mathbf{U}$ on the values of t_i makes explicit solutions to the design problem difficult to obtain even in the simplest cases. By keeping T fixed and letting $N \to \infty$, Sacks and Ylvisaker (1966, 1968) were able to develop an asymptotic theory for (1.1) and other models. Unfortunately their theory yields explicit solutions only in a few special cases; moreover, the solutions seem to depend strongly on the assumed form of dependence.

Here interest will be in situations where the dependence of the observations is not known too precisely and at least initially is assumed to be negligible. This leads to a consideration of a different type of asymptotics in which one not only lets $N \to \infty$ but also permits ρ to depend on N and "be close" to the correlation function for independent errors if N is large. Specifically it is assumed that for N observations, the correlation function is given by

$$\text{(1.9)} \qquad \rho_N(t) = \rho_1(Nt),$$

where, at the very least, $\rho_1(t) \to 0$ as $t \to \infty$.

Such asymptotics can be justified on two grounds.

(i) The variance-covariance matrices of the least squares estimates and a fortiori the best linear unbiased estimators are of the order of $1/N$ as in the case of independence rather than 0(1) as in the Sacks-Ylvisaker theory. Thus the independence and dependence situations are comparable and permit one to think in the usual robustness terms of sacrificing some efficiency of design in the independence situation in exchange for a great gain in efficiency in the dependence situation.

(ii) If the f_i are powers, asymptotic results for varying ρ as in (1.9) can be translated into asymptotic results for $\rho = \rho_1$ and varying $T = T_N = NT$. This seems reasonable since, typically, as the number of observations that can be taken is increased the interval of time during which observations can be taken increases also; see Comment 3 following Theorem 2.1.

The initial assumption that ρ is negligible suggests that instead of considering the best linear unbiased estimators, designs should be found for which the least squares estimates perform well. That is, we should consider the asymptotic theory of the variance-covariance matrix of the least squares estimates, $\sigma^2\tilde{\boldsymbol{\Sigma}}$, where

$$\tilde{\boldsymbol{\Sigma}} = (\mathbf{F}^T\mathbf{F})^{-1}(\mathbf{F}^T\mathbf{U}\mathbf{F})(\mathbf{F}^T\mathbf{F})^{-1}. \tag{1.10}$$

The limiting behavior of $\tilde{\boldsymbol{\Sigma}}$ for smooth designs is given in Section 2. In the remainder of the paper, attention is restricted to the simplest cases of model (1.1), i.e., estimation of location and regression through the origin,

$$p = 1,$$

and

$$f_1(t) = 1,$$

$$f_1(t) = t,$$

and simple linear regression.

In Section 3, the formulae derived in Section 2 are used to solve the asymptotic versions of the optimality problem, namely

$$0\text{: Minimize } N\tilde{\boldsymbol{\Sigma}} \text{ over designs } (t_1, \ldots, t_N);$$

and a generalization appropriate to a robust formulation,

$$0' : \text{Minimize } N\tilde{\boldsymbol{\Sigma}} \text{ over designs } (t_1, \ldots, t_N)$$

$$\text{subject to } N(\mathbf{F}^T\mathbf{F})^{-1} \leqslant \lambda T^{-2}\mathbf{I} \qquad (\lambda > 1).$$

The version $0'$ formalizes the idea that a reasonable but suboptimal performance should be required if $\rho = 0$ and subject to that requirement one should try to do as well as possible in the presence of dependence.

The solutions obtained in Section 3 are not completely explicit and depend on the form of ρ and on γ. In Section 4 a further asymptotic analysis is given which to some extent reverses the asymptotics of previous sections by letting $T \to 0$. The pleasing result is obtained that the classical equispaced design $t_i = \{2(i-1)/(N-1) - 1\}T (i = 1, \ldots, N)$ is asymptotically a solution to 0, and to $0'$ if $\lambda \geqslant 3$, under weak conditions on ρ and γ.

In a companion paper, Bickel, Herzberg and Schilling (1977), some exact calculations are given for $\rho(t) = e^{-\lambda|t|}$ which enable one to judge the adequacy of the asymptotics.

2. Asymptotic theory. Consider sequences of designs

$$\{t_{1N} \leqslant \cdots \leqslant t_{NN}\} \qquad N \geqslant 1$$

defined by a continuous nondecreasing function

$$a : [0, 1] \to [-T, T]$$

through

$$t_{iN} = a\left(\frac{i-1}{N-1}\right) \qquad i = 1, \ldots, N. \tag{2.1}$$

Clearly, any design can be embedded in such a sequence and any such sequence determines $a(\cdot)$ uniquely. There is usually a 'natural' choice of $a(\cdot)$. Thus, the equispaced design corresponds to $a(t) = (2t - 1)T$; the design having all observations at T has $a \equiv T$. There is a simple correspondence between this notion and the more usual design measures. If ξ_N denotes the design measure of the Nth design, then the ξ_N tend weakly to ξ, where $a(\cdot)$ is the inverse of the distribution function of ξ, i.e.,

$$\xi(A) = \lambda\{t : a(t) \in A\}, \tag{2.2}$$

where $\lambda(\cdot)$ is the Lebesgue measure on [0, 1].

The following regularity conditions for the validity of our asymptotic formulae are required. For convenience from now on we write ρ for ρ_1.

A. *Conditions on a(t).* The function $a(t)$ is twice differentiable on (0, 1) and there exists a value $M < \infty$ such that

$$M^{-1} \leqslant a'(t) \leqslant M, \tag{2.3}$$

$$|a''(t)| \leqslant M. \tag{2.4}$$

R. *Conditions on* $\rho(t)$. (i) The function $\rho(t)$ is differentiable on $(0, \infty)$ and there exists $M < \infty$ such that

$$|\rho'(t)| \leqslant M. \tag{2.5}$$

(ii) Moreover, $\rho'(t) \leqslant 0$ for t sufficiently large and

$$\int_0^\infty |\rho(s)|ds < \infty. \tag{2.6}$$

Note that (ii) implies that $\rho(t)$ is nonnegative for t sufficiently large.

F. *Conditions on* $f_k(t)$. The functions $f_k(t)$ obey a first order Lipschitz condition,

$$|f_k(t) - f_k(s)| \leqslant M|t - s| \tag{2.7}$$

for all s, $t (k = 1, \ldots, p)$. Without loss of generality suppose also

$$|f_k(t)| \leqslant M \tag{2.8}$$

for all k.

Note that (ii) guarantees that the function Q given by

$$Q(t) = \Sigma_{j=1}^{\infty}\rho(jt) \tag{2.9}$$

is well defined and finite for all $t > 0$. Here is the key result.

LEMMA 2.1. *Suppose* $a(\cdot)$, $f_k(\cdot)$ *and* $\rho(\cdot)$ *satisfy* (2.3)–(2.8). *Then*

$$N^{-1}\Sigma_{i\neq j}f_k(t_{iN})f_l(t_{jN})\rho\{N(t_{jN} - t_{iN})\} = 2\int_0^1 f_k\{a(t)\}f_l\{a(t)\}Q\{a'(t)\}dt + o(1) \tag{2.10}$$

provided that the first term on the right-hand side of (2.10) *is finite.*

PROOF. Write

$$\begin{aligned} N^{-1}\Sigma_{i\neq j}f_k(t_{iN})f_l(t_{jN})\rho\{N(t_{jN} - t_{iN})\} &= 2\big[N^{-1}\Sigma_{i=1}^{N}f_k(t_{iN})f_l(t_{iN})\Sigma_{j=i+1}^{N}\rho\{N(t_{jN} - t_{iN})\} \\ &\quad + N^{-1}\Sigma_{i=1}^{N}f_k(t_{iN})\Sigma_{j=i+1}^{N}\{f_l(t_{jN}) - f_l(t_{iN})\}\rho\{N(t_{jN} - t_{iN})\}\big]. \end{aligned} \tag{2.11}$$

Let $r_N \to \infty$. Then, from (2.8),

$$|\Sigma_{j>i+r_N}\{f_l(t_{jN}) - f_l(t_{iN})\}\rho\{N(t_{jN} - t_{iN})\}| \leqslant MT\Sigma_{j>i+r_N}|\rho|\{N(t_{jN} - t_{iN})\}. \tag{2.12}$$

Since ρ is eventually nonincreasing and (2.3) holds, the right-hand side of (2.12) is

$$0\big[\Sigma_{j>i+r_N}\rho\{M^{-1}(j - i)\}\big].$$

But

$$\Sigma_{k>r_N}\rho(M^{-1}k) = o(1)$$

by the convergence in (2.9). Similarly

$$N^{-1}\Sigma_{i=1}^{N}f_k(t_{iN})f_l(t_{iN})\Sigma_{j>i+r_N}\rho\{N(t_{jN} - t_{iN})\} = o(1). \tag{2.13}$$

Write N_1 for $N - 1$, i_N for $(i - 1)/(N - 1)$.

On the other hand, by (2.4) and (2.5),

$$\begin{aligned} \rho\{N(t_{jN} - t_{iN})\} &= \rho\left\{a'(i_N)(j - i) + \frac{\theta(j - i)^2}{N_1}\right\} \\ &= \rho\{a'(i_N)(j - i)\} + \theta'\frac{(j - i)^2}{N_1}, \end{aligned} \tag{2.14}$$

where $|\theta| \leqslant \frac{1}{2}M$ and $|\theta'| \leqslant \frac{1}{2}M^2$.

From (2.7) and (2.8),

$$|N^{-1}\Sigma_{i=1}^{N}f_k(t_{iN})\Sigma_{j=i+1}^{i+r_N}\{f_l(t_{jN}) - f_l(t_{iN})\}\rho\{N(t_{jN} - t_{iN})\}| \leqslant M^2\max_i\Sigma_{j=i+1}^{i+r_N}(t_{jN} - t_{iN})|\rho\{N(t_{jN} - t_{iN})\}|. \tag{2.15}$$

But

$$t_{jN} - t_{iN} \leqslant \frac{M}{N_1}(j - i)$$

and by (2.14) the right-hand side of (2.15) is bounded by

(2.16) $$M^3 \max_i \left[N_1^{-1} \Sigma_{j=i+1}^{i+r_N} (j-i) \rho\{a'(i_N)(j-i)\} + N^{-2} \Sigma_{j=i+1}^{i+r_N} (j-i)^3 \right].$$

The first term in (2.16) is $0(r_N/N)$ in view of (2.3) and (2.6) while the second term is $0(r_N^4/N^2)$.

Similarly, from (2.14),

(2.17) $$N^{-1} \Sigma_{i=1}^N f_k(t_{iN}) f_l(t_{iN}) \Sigma_{j=i+1}^{i+r_N} \rho\{N(t_{jN} - t_{iN})\}$$

$$= N^{-1} \Sigma_{i=1}^N f_k(t_{iN}) f_l(t_{iN}) \Sigma_{j=i+1}^{i+r_N} \rho\{a'(i_N)(j-i)\} + 0\left(\frac{r_N^3}{N}\right)$$

$$= N^{-1} \Sigma_{i=1}^N f_k(t_{iN}) f_l(t_{iN}) Q\{a'(i_N)\} + 0\left[\Sigma_{j=r_N+1}^{\infty} |\rho\{a'(i_N)j\}|\right] + 0\left(\frac{r_N^3}{N}\right).$$

If we let $r_N \to \infty$, so that $r_N = o(N^{\frac{1}{3}})$, we can conclude from (2.10), (2.12), (2.13), (2.16) and (2.17) that

(2.18) $$N^{-1} \Sigma_{i \neq j} f_k(t_{iN}) f_l(t_{jN}) \rho\{N(t_{jN} - t_{iN})\}$$

$$= 2N^{-1} \Sigma_{i=1}^N f_k(t_{iN}) f_l(t_{iN}) Q\{a'(i_N)\}$$

$$= 2\int_0^1 f_k(t) f_l(t) Q\{a'(t)\} dt + o(1)$$

and the result follows.

The main result of the section follows as an immediate consequence of the lemma.

THEOREM 2.1. *Suppose* $a(\cdot)$, *the* $f_k(\cdot)$ *and* $\rho(\cdot)$ *satisfy* (2.3)–(2.8) *and the* $f_k(\cdot)$ *are not linearly related on a set of positive Lebesgue measure. Suppose also that the elements of the* $p \times p$ *matrix* $\mathbf{R}(a)$ *exist and are finite, where*

$$\mathbf{R}(a) = \left\{ \int_0^1 f_i\{a(t)\} f_j\{a(t)\} Q\{a'(t)\} dt \right\} \qquad i, j = 1, \ldots, p.$$

Then the matrix $\tilde{\boldsymbol{\Sigma}}$ *given by* (1.10) *is well defined for N sufficiently large and*

(2.19) $$N\tilde{\boldsymbol{\Sigma}} \to \mathbf{U}(a) = \mathbf{W}^{-1}(a)\{\mathbf{I} + 2\gamma \mathbf{R}(a) \mathbf{W}^{-1}(a)\},$$

where

(2.20) $$W(a) = \left[\int_0^1 f_i\{a(t)\} f_j\{a(t)\} dt \right] \qquad i, j = 1, \ldots, p$$

and $\mathbf{I}$ *is the* $p \times p$ *identity matrix.*

PROOF. Write

(2.21) $$\mathbf{F}^T \mathbf{U} \mathbf{F} = \left[\gamma \Sigma_{i \neq j} f_k(t_{iN}) f_l(t_{jN}) \rho\{N(t_{jN} - t_{iN})\} + \Sigma_{i=1}^N f_k(t_{iN}) f_l(t_{iN}) \right] \qquad k, l = 1, \ldots, p$$

and apply Lemma 2.1 and standard limiting arguments.

COMMENTS. 1. The conditions on $a(\cdot)$ imply that ξ defined by (2.2) is absolutely continuous with continuous density ξ'. We can express **W** and **R** and hence **U** in terms of ξ simply by

$$\mathbf{W}(a) = \{\textstyle\int_{-T}^{T} f_k(t) f_l(t) d\xi(t)\} \qquad k, l = 1, \ldots, p, \tag{2.22}$$

$$\mathbf{R}(a) = \left[\textstyle\int_{-T}^{T} f_k(t) f_l(t) Q\left\{\frac{1}{\xi'(t)}\right\} d\xi(t)\right] \qquad k, l = 1, \ldots, p. \tag{2.23}$$

2. If **R** is not well defined, but the other conditions of the theorem are satisfied $N\tilde{\Sigma}$ may not converge. We do not know what happens if $a(\cdot)$ does not satisfy conditions (2.3) and (2.4). The expression (2.19) for $\mathbf{R}(a)$ can still make sense if we interpret $a'(\cdot)$ as the a.e. well-defined derivative of the continuous nondecreasing function $a(\cdot)$ and $Q(0) = \infty$. In some cases it is possible to check that $\mathbf{U}(a)$ is still the limit of $N\tilde{\Sigma}$. For instance, if $p = 1$ and $f_1(t) = t$, $\mathbf{R}(a) = \infty$ if $a(\cdot)$ is constant and nonzero on an interval. If $\rho > 0$, it is easy to see that for such $a(\cdot)$, $N\tilde{\Sigma} \to \infty$. An example of such an $a(\cdot)$ is the optimal design for linear regression through the origin when the errors are independent.

3. Suppose that instead of letting $\rho(\cdot)$ vary with N we fix $\rho(\cdot)$ but permit observations to be taken on the interval $[-TN, NT]$ at design points $Na(i_N)$ $(i = 1, \ldots, N)$, where $a(\cdot)$ is as in the statement of the theorem. If $f_k(t) = t^{k-1}$ and $\hat{\beta}_k$ $(k = 1, \ldots, p)$ are the least squares estimates, then $\tilde{\Sigma}$ given by (1.10) is just the variance covariance matrix of $\hat{\beta}_k N^k$ $(k = 0, 1, \ldots, p)$. Thus asymptotic comparisons among different designs in this formulation are also made on the basis of the matrices $\mathbf{U}(a)$.

4. For the special case considered in paper II, i.e., $p = 1, f_1(t) = t$, $\rho(t) = e^{-\lambda|t|}$, it can be shown that $N\Sigma$ and $N\tilde{\Sigma}$ have asymptotically the same limit. Grenander and Rosenblatt (1957) show the equivalence for the equispaced design, f_k as in comment 3 and a wide class of ρ's. We do not know how generally the equivalence holds in our context.

5. Results related to Theorem 2.1 may be found in Grenander and Rosenblatt (1957) and Brillinger (1973).

3. Optimal designs for p = 1 and linear regression. Suppose that $p = 1$. Then, for the estimates corresponding to an $a(\cdot)$ satisfying (2.3) and (2.4), the standardized variance, $N\tilde{\Sigma}$, tends to

(3.1)

$$\mathbf{U}(a) = [\textstyle\int_0^1 f_1^2\{a(t)\}dt]^{-1}\left(1 + 2\gamma \int_0^1 f_1^2\{a(t)\} Q\{a'(t)\} dt [\int_0^1 f_1^2\{a(t)\}dt]^{-1}\right).$$

Although $\mathbf{U}(a)$ is only valid for $a(\cdot)$ satisfying (2.3) and (2.4), we shall ignore these restrictions and consider the general minimization problems:

O: Minimize $\mathbf{U}(a)$ over all $a(\cdot)$ which are the inverses of absolutely continuous probability distributions on $[-T, T]$.

O′: Minimize $\mathbf{U}(a)$ over all $a(\cdot)$ as in 0 which in addition satisfy $\int_0^1 f_1^2\{a(t)\}dt \geqslant \lambda$.

We shall show that these problems have solutions and at least partially calculate these solutions under a strong condition on Q.

R(iii). The function Q is strictly convex on $(0, \infty)$.

Note that R(iii) holds if ρ is strictly convex. We then define, for $t > 0$,

$$H(t) = Q(t) - tQ'(t). \tag{3.2}$$

Now $Q(0+) = \infty$, Q is strictly convex, and by R(ii), $Q(t)$ is decreasing for t sufficiently large,

$$Q(\infty) = 0. \tag{3.3}$$

We conclude that $Q(t) \geqslant 0$ and is decreasing on $(0, \infty)$. Therefore,

$$H(t) > 0 \qquad (t > 0), \tag{3.4}$$
$$H(0_+) = \infty, \qquad H(\infty) = 0.$$

Moreover, since

$$H'(t) = -tQ''(t), \tag{3.5}$$

we see that H is strictly decreasing on $(0, \infty)$. By (3.4) and (3.5), H has a well defined inverse $H^{-1}[0, \infty] \to [0, \infty]$ where, of course, $H^{-1}(0) = \infty$, $H^{-1}(\infty) = 0$.

We require a condition weaker than R(iv) of Section 4, i.e.,

$$0 < \varepsilon \leqslant \liminf_{t \to 0} t^2 |Q'(t)| \leqslant \limsup_{t \to 0} t^2 |Q'(t)| \leqslant \varepsilon^{-1} < \infty, \tag{3.6}$$

i.e., $Q'(t) = \Omega(t^{-2})$, as $t \to 0$.

Define functions $q(\cdot, \mu, \tau)$ by

$$q(x, \mu, \tau) = \left(H^{-1}\left[\mu\{1 - \tau f_1^{-2}(x)\}\right]\right)^{-1} \quad \text{if } \mu\{1 - \tau f_1^{-2}(x)\} \geqslant 0, \quad |x| \leqslant T, \tag{3.7}$$
$$= 0 \qquad \text{otherwise}.$$

THEOREM 3.1. (i) *If* $f_1(\cdot)$ *is continuous on* $[-T, T]$, R(iii) *and* (3.6) *hold then solutions to problems O and O′ exist.*

(ii) *The optimal* $a^*(\cdot)$ *correspond to probability densities* $p^*(\cdot)$ *which either*

(a) *are of the form* $q(\cdot, \mu, \tau)$ *for suitable* μ^*, τ^*; *or*

(b) *are uniform distributions on sets of constancy of* $f_1^2(\cdot)$; *or*

(c) *are constant on the set* $\{f_1^2(t) \neq 0\}$ *and arbitrary otherwise.*

(iii) *The solution to O is always of the form* $q(\cdot, \mu^*, \tau^*)$, *where* μ^*, τ^* *satisfy the equations*

$$\int_{-T}^{T} q(t, \mu^*, \tau^*)\,dt = 1, \tag{3.8}$$

$$\frac{2\int_{-T}^{T} Q\left(\frac{1}{q}\right) f_1^2 q\,dt}{\int_{-T}^{T} f_1^2 q\,dt} = \mu^* - \frac{1}{2\gamma}. \tag{3.9}$$

The rather technical proof of this result relies on solving the associated Lagrange problem. The proof is given in the Appendix.

Clearly $\mu^* > 0$. Solutions $q(\cdot, \mu^*, \tau^*)$ of O' as well as O have $\mu^* > 0$ and $\tau^* \geqslant 0$ under a simple condition stated at the end of the Appendix.

We are unable to show that (3.8) and (3.9) determine p^* uniquely except in the trivial case $f_1(t) \equiv 1$. However, we are able to use these necessary conditions to conclude that the uniform design is asymptotically optimal as $T \to 0$ if $f_1(t) = t$.

EXAMPLE 1. *Location.* If $f_1(t) \equiv 1$, all $q(\cdot, \mu, \tau)$ are constant on $[-T, T]$. Problems 0 and 0′ coincide and the optimal p^* is, necessarily, the uniform design

$$p^*(t) = \frac{1}{2T}(|t| \leqslant T), \qquad = 0 \quad \text{otherwise.} \tag{3.10}$$

EXAMPLE 2. *Regression through the origin.* If $f_1(t) = t$, formula (3.7) leads to two families of functions of differing shapes:

I. $\mu > 0, 0 \leqslant \tau < T^{\frac{1}{2}}$,

$$\begin{aligned} q(t, \mu, \tau) &= 0 && (|t| \leqslant \tau^{\frac{1}{2}}), \\ &= [H^{-1}\{\mu(1 - \tau t^{-2})\}]^{-1} && (\tau^{\frac{1}{2}} \leqslant |t| \leqslant T), \\ &= 0 && (|t| > T). \end{aligned}$$

II. $\mu > 0, \sigma = -\tau \geqslant 0$,

$$\begin{aligned} q(t, \mu, -\sigma) &= [H^{-1}\{\mu(1 + \sigma t^{-2})\}]^{-1} && (|t| \leqslant T), \\ &= 0 && \text{otherwise.} \end{aligned}$$

But (3.6) implies that all functions of type II have $\int_{-T}^{T} q(t, \mu, \tau) = \infty$ and need not be considered. Since H^{-1} is decreasing, for fixed τ, $\int_{-T}^{T} q(t, \mu, \tau)dt$ is increasing in μ from 0 to ∞ as μ ranges from 0 to ∞. Thus for each $0 \leqslant \tau < T^{\frac{1}{2}}$, there is a unique $\mu(\tau)$ such that $q(\cdot, \mu(\tau), \tau)$ satisfies (3.8). A typical member of the resulting families of densities is shown in Figure 1. The borderline case $\tau = 0$ of course has $\mu = H(2T)$ and is again the uniform density.

Moreover, if $\tau_1 < \tau_2$, $q(t, \mu(\tau_1), \tau_1)$ and $q(t, \mu(\tau_2), \tau_2)$ cross exactly once for $t > 0$ from + to −. Therefore, $\int_{-T}^{T} t^2 q(t, \mu(\tau), \tau)dt$ is an increasing function of τ from $\frac{1}{3}T^2$ when $\tau = 0$ to T^2 as $\tau \to T^{\frac{1}{2}}$. Thus, we can reparameterize $q(\cdot, \mu(\tau), \tau)$ continuously as $\pi(\cdot, \eta)$ say, where $\eta = \int_{-T}^{T} t^2 q(t, \mu(\tau), \tau)dt$. Since $f_1^2(t)$ has no sets of constancy of positive measure, the solution to O' is obtained by minimizing

$$G(\eta) = \frac{1}{\eta}\left[1 + 2\gamma \int t^2 Q\left\{\frac{1}{\pi(t, \eta)}\right\}\pi(t, \eta)dt\right]$$

over $\eta \geqslant \lambda$. It is easy to see in this context that (3.9) corresponds to the equation, $G'(\eta^*) = 0$, where η^* corresponds to μ^* *and* τ^*. If η^* is unique, it follows that the solution to O' is η^* if $\eta^* \geqslant \lambda$ and λ otherwise since $G(T^2) = \infty$.

EXAMPLE 3. *Linear regression.* Consider $p = 2, f_1(t) = 1, f_2(t) = t$. Any optimization problem here must be formulated in terms of a function of the matrix **U**

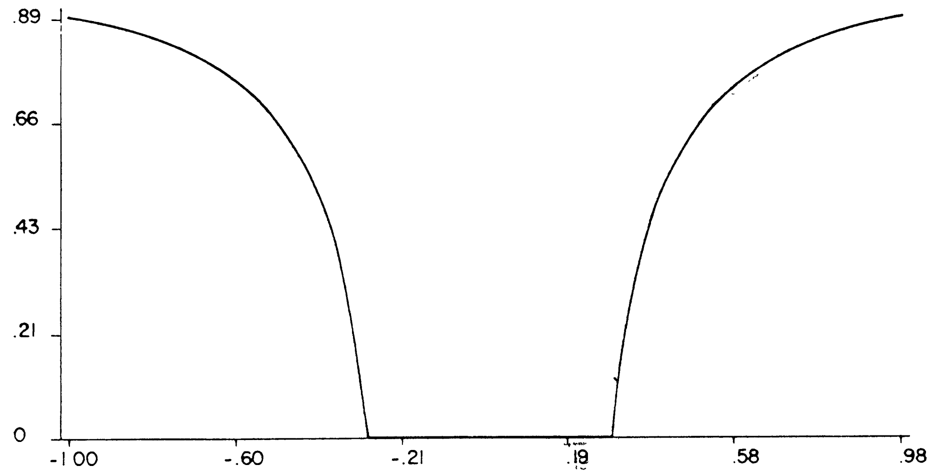

FIG. 1. *The density* q (•, 3.5, 0.0891 *optimal for* $T = 1, \gamma = .5, \rho(t) = e^{-0.483|t|}$.

given in (2.19). However, the solution is much simpler if attention is restricted to symmetric designs such that $a(t) = -a(1-t)$ $(0 < t < 1)$. Then $\mathbf{U}$ becomes diagonal with diagonal elements

$$u_{11}(a) = 1 + 2\gamma \int_0^1 Q\{a'(t)\}\,dt,$$

$$u_{22}(a) = \left\{\int_0^1 a^2(t)\,dt\right\}^{-1}\left[1 + 2\gamma \frac{\int_0^1 Q\{a'(t)\}a^2(t)\,dt}{\int_0^1 a^2(t)\,dt}\right],$$

the functionals being considered in Examples 1 and 2. It is easy to see that minimizing any convex function of $(u_{11}(a), u_{22}(a))$ over the set of symmetric $a(\cdot)$'s leads to a solution in the family $q(\cdot, \mu(\tau), \tau)$.

NOTE. In the ordinary theory of experimental design with independent errors, the assumption of normality does not affect the choice of the optimal design at least asymptotically; see, for example, Cox and Hinkley (1968). We believe this is true in the cases we consider also but the matter clearly needs further study.

4. Asymptotic optimality as $\mathbf{T \to 0}$. We have seen that the uniform design is optimal for location. It is typically not optimal for regression through the origin. For example, if $\rho(t) = e^{-\lambda|t|}$, (3.9) is not satisfied for any value of γ and $T > 0$. However, it is asymptotically optimal as $T \to 0+$ under some conditions on the local behavior of $f_1(\cdot)$ and $Q(\cdot)$ at 0. For simplicity we give details for the case $f_1(t) = t$ only. Assume:

R(iv). As $t \to 0$, for c, d finite

(4.1) $$tQ(t) = c + 0(1),$$

(4.2) $$\{tQ(t)\}' = d + 0(1).$$

Let U_T denote the minimum value of $U(a)$.

THEOREM 4.1. *If* R(iii) *and* R(iv) *hold and* $f_1(t) = t$, *then*

$$U_T = U((2t-1)T) + o(T^{-2}) \tag{4.3}$$
$$= 3T^{-2}(\gamma c T^{-1} + 1 + 2\gamma d) + o(T^{-2}).$$

LEMMA 4.1. *Under* (4.1) *and* (4.2), *as* $t \to 0$,

$$Q(t) = ct^{-1} + d + o(1), \tag{4.4}$$

and, as $x \to \infty$,

$$\{H^{-1}(x)\}^{-1} = \frac{x-d}{2c} + o(1), \tag{4.5}$$

$$Q\{H^{-1}(x)\} = \frac{1}{2}(x+d) + o(1). \tag{4.6}$$

PROOF. Now (4.4) is equivalent to

$$tQ(t) = c + td + o(t),$$

an elementary consequence of (4.1) and (4.2). Combining (4.2) and (4.4), we get as $t \to 0$

$$H(t) = \frac{2c}{t} + d + o(1) \tag{4.7}$$

and (4.5) and (4.6) follow.

PROOF OF THEOREM 4.1. Let μ_T and τ_T be the optimal values of μ^* and τ^* guaranteed by Theorem 3.1(iii). We argue for (4.3) under the assumption $\tau_T \geqslant 0$. A similar condition works for sequences $\tau_T \leqslant 0$. For convenience in computation we drop the T subscripts on μ and τ. Also we define $\rho = \tau^{\frac{1}{2}}T^{-1}$. By (3.7) and (3.8)

$$1 = 2\int_{\tau^{\frac{1}{2}}}^{T}[H^{-1}\{\mu(1-\tau t^{-2})\}]^{-1}dt \tag{4.8}$$
$$= 2T\int_{\rho}^{1}[H^{-1}\{\mu(1-\rho^2 w^{-2})\}]^{-1}dw$$
$$\leqslant 2T\{H^{-1}(\mu)\}^{-1}. \tag{4.9}$$

By (4.9), $\mu \to \infty$. Upon applying (4.5) to (4.8) we obtain,

$$1 = \frac{T}{c}\int_{\rho}^{1}\{\mu(1-\rho^2 w^{-2}) - d + \varepsilon_T(w)\}dw \tag{4.10}$$

where $|\varepsilon_T|$ is bounded and $\sup[|\varepsilon_T(w)| : w \geqslant \rho\{1 - M(T)\mu^{-1}\}^{-\frac{1}{2}}] \to 0$ for any sequence $M(T) \to \infty$, $M(T) \leqslant \mu$. Therefore (4.10) becomes after some algebra,

$$\frac{T}{c}(1-\rho)\{\mu(1-\rho) - d\} = 1 + o(T). \tag{4.11}$$

Similarly,

$$(4.12)\qquad 2\int_{\tau^{\frac{1}{2}}}^{T} t^2\left[H^{-1}\{\mu(1-\tau t^{-2})\}\right]^{-1}dt$$
$$= 2T^3\int_{\rho}^{1} w^2\left[H^{-1}\{\mu(1-\rho^2 w^{-2})\}\right]^{-1}dw$$
$$= \tfrac{1}{3}c^{-1}T^3(1-\rho)\{u(1-\rho)(1+2\rho)-d\} + o(T^3) + 0(\rho T^3).$$

Also, since $Q\{H^{-1}(x)\}\{H^{-1}(x)\}^{-1} = \frac{1}{4}c^{-1}x^2 + o(x)$ as $x \to \infty$,

$$(4.13)\qquad 2\int_{\tau^{\frac{1}{2}}}^{T} t^2 Q\left[H^{-1}\{\mu(1-\tau t^{-2})\}\right]\left[H^{-1}\{\mu(1-\tau t^{-2})\}\right]^{-1}dt$$
$$= \tfrac{1}{6}c^{-1}T^3\left[\int_{\rho}^{1}\mu^2(w^2-2\rho^2+\rho^4 w^{-2})dw + o\left\{\int_{\rho}^{1}\mu(w^2-\rho^2)dw\right\} + o(1)\right]$$
$$= \tfrac{1}{6}c^{-1}T^3\mu^2\{(1-\rho)^3(3\rho+1) + o(\mu^{-1})\}.$$

If we substitute (4.12) and (4.13) in equation (3.9), we get

$$(4.14)\qquad \mu\left\{(1-\rho)^2(3\rho+1) + o\left(\frac{1}{\mu}\right)\right\}\left\{(1-\rho)(2\rho+1) - \frac{d}{\mu} + 0\left(\frac{\rho}{\mu}\right) + o\left(\frac{1}{\mu}\right)\right\}^{-1}$$
$$= \mu\{1-(2\gamma\mu)^{-1}\}.$$

Thus $\rho \to 0$, and, collecting terms,

$$(4.15)\qquad \rho^2 = \frac{1}{3\mu}\left(\frac{1}{2\gamma} + d\right) + o\left(\frac{1}{\mu}\right).$$

Now from (3.9)

$$(4.16)\qquad U_T = \left(2\int_{\tau^{\frac{1}{2}}}^{T} t^2\left[H^{-1}\{\mu(1-\tau t^{-2})\}\right]^{-1}dt\right)^{-1}\left(\gamma\mu + \tfrac{1}{2}\right)$$

$$(4.17)\qquad = \frac{3cT^{-3}}{\mu}\left(\gamma\mu + \tfrac{1}{2}\right)\left\{1 + 3\rho^2 + \frac{d}{\mu} + o\left(\frac{1}{\mu}\right) + o(\rho^2)\right\}$$

by (4.12).

Simplifying further and using (4.15) we obtain

$$(4.18)\qquad U_T = 3\left\{\gamma T^{-3}c + \frac{cT^{-3}}{\mu}(1+2\gamma d)\right\} + o(T^{-3}\mu^{-1}).$$

Since $\rho \to 0$, (4.14) implies that

$$(4.19)\qquad \mu = cT^{-1} + o(T^{-1})$$

and hence, from (4.18),

$$(4.20)\qquad U_T = 3\{\gamma c T^{-3} + T^{-2}(1+2\gamma d)\} + o(T^{-2}).$$

On the other hand, by (3.1),

$$U(2t-T) = 3T^{-2}\{1 + 2\gamma Q(2T)\}$$
$$= 3T^{-2}\left[1 + 2\gamma\{c(2T)^{-1} + d\}\right] + o(T^{-2})$$
$$= U_T + o(T^{-2}).$$

The theorem follows.

NOTE. Asymptotic optimality of the uniform design holds under more general conditions on $f_1(\cdot)$ as Example 1 would suggest. All that is needed is

F(ii). $f_1(t) = at^b + o(t^b)$

as $t \to 0$, where $a \neq 0$ and b is a nonnegative integer.

APPENDIX

Throughout the Appendix, it is assumed that $f_1(\cdot)$ is continuous on $[-T, T]$, R(iii) and (3.6) hold. In order to prove Theorem 3.1 we need to study some Lagrange problems.

Let

(1) $$P = \{p : p \geqslant 0, \int\{f_1^2(t)p^2(t) + p(t)\}dt < \infty\}.$$

For p for which the integral involved is well defined, let

(2) $$V(p) = \int_{-T}^{T}\left(f_1^2(t)\left[Q\left\{\frac{1}{p(t)}\right\} - \mu\right] + \mu\tau\right)p(t)dt.$$

Conditions (3.6) and (3.3) imply that, as $t \to 0$,

(3) $$Q(t) = \Omega(t^{-1}).$$

Therefore, $V(p)$ is finite for $p \in P$ and conversely $V(p)$ finite implies $p \in P$.

Write q for $q(\cdot, \mu, \tau)$ defined by (3.7).

LEMMA 1. *The function* $V(q)$ *is always well defined and*

(4) $$V(q) = \inf\{V(p) : p \in P\}.$$

If $\inf\{V(p) : p \in P\} > -\infty$, *then* $q \in P$. *Moreover, if* (4) *holds and* $V(q_n) \to V(q)$, $q_n \in P$, *then*

(5) $$\int_{-T}^{T} f_1^2(t)\{q_n(t, \mu, \tau) - q(t, \mu, \tau)\}^2 dt \to 0.$$

PROOF. If $p_0, p \in P$, let $g(\varepsilon) = V(p_\varepsilon)$, where $p_\varepsilon = p_0 + \varepsilon(p - p_0)$. Since, by (3.6), $H(t) = \Omega(t^{-1})$, we can apply dominated convergence to conclude that $g(\cdot)$ is differentiable and

(6) $$g'(\varepsilon) = \int_{-T}^{T}\left(f_1^2(t)\left[H\left\{\frac{1}{p_\varepsilon(t)}\right\} - \mu\right] + \mu\tau\right)\{p(t) - p_0(t)\}dt.$$

Moreover, $g(\cdot)$ is convex since $H(\cdot)$ is decreasing and

(7)
$$g'(\varepsilon + h) - g'(\varepsilon) = h\int_{-T}^{T}\left\{H\left(\frac{1}{p_{\varepsilon+h}}\right) - H\left(\frac{1}{p_\varepsilon}\right)\right\}(p_{\varepsilon+h} - p_\varepsilon)^{-1}(p - p_0)^2 \geqslant 0.$$

Let $p \in P$. Define

$$p^M = p(\cdot) \quad \text{if} \quad q(\cdot, \mu, \tau) \geqslant M, \qquad q \geqslant p$$
$$= q(\cdot, \mu, \tau) \text{ otherwise.}$$

Then $p^M \in P$. Let $p^M = p_0$. Then

$$(8) \qquad g'(0) = \int_{[p^M=q]}\left(f_1^2(t)\left[H\left\{\frac{1}{q(t,\mu,\tau)}\right\} - \mu\right] + \mu\tau\right)\{p(t) - q(t,\mu,\tau)\}dt.$$

Since $p \geqslant 0$ and $q(\cdot, \mu, \tau)$ is given by (3.7), the integral is always nonnegative. Since $g(\cdot)$ is convex, we conclude that

$$(9) \qquad V(p^M) \leqslant V(p).$$

Since $p^M \uparrow q$ as $M \to \infty$, we can apply monotone convergence to conclude that

$$(10) \qquad V(p^M) \downarrow V(q),$$

where the right-hand side must be well defined and (4) follows.

Note that the integrand in $V(p_\varepsilon)$ can be expanded in a Taylor's series with Cauchy's form of the remainder. Using finiteness of $g'(0)$ and Fubini's theorem, we obtain

$$(11) \qquad g(\varepsilon) = g(0) + \varepsilon g'(0) + \varepsilon^2\int_0^1(1-\lambda)\int f_1^2(x)\left[-p_{\varepsilon\lambda}^{-2}(x)H^{-1}\{p_{\varepsilon\lambda}^{-1}(x)\}\right](p-p_0)^2(x)\,dx\,d\lambda.$$

For $t > 0$, let $0 < M(t) = \inf\{-x^{-2}H'(x^{-1}): x \geqslant t\}$. Since $p - p_0 \geqslant \delta$ implies $p_\lambda \geqslant \lambda\delta$, while $p - p_0 \leqslant 0$ implies $p_\lambda \geqslant (1-\lambda)p_0$, the last integral in (11) is greater than or equal to

$$\tfrac{1}{2}\varepsilon^2 M\left(\tfrac{1}{2}\delta\right)\int_{A_\delta} f_1^2(x)(p-p_0)^2dx,$$

where $A_\delta = \{p - p_0 \geqslant \delta \text{ or } p \leqslant p_0 \text{ and } p_0 \geqslant \delta\}$. Now put $\varepsilon = 1$, $p_0 = q$, $p = q_n$ and obtain

$$(12) \qquad V(q_n) - V(q) \geqslant \int_{A_\delta} f_1^2(x)(q_n - q)^2dx.$$

Since we also have

$$(13) \qquad \int_{A_\delta} f_1^2(x)(q_n - q)^2dx \leqslant 2T\delta^2 \max f_1^2(x),$$

(5) follows.

LEMMA 2. *Suppose $A \subset [-T, T]$ and $m(A) > 0$, where m denotes Lebesgue measure. Let $p_A = (1/m(A))I_A$, the uniform density function on A. Then, for any $c > 0$,*

$$(14) \qquad \int_A Q(cp_A^{-1})p_A = \min\left\{\int_A Q(cp^{-1})p : p \geqslant 0, \int_A p = 1\right\}.$$

PROOF. Arguing as in Lemma 1, we can show that p_A minimizes (for suitable μ) $\int_A Q(cp^{-1})p - \mu\int p$ for all $p \geqslant 0$. Since p_A satisfies the condition $\int_A p_A = 1$, the result follows.

We now relate Problem O to our Lagrange lemmata.

Let

$$S = \left\{\left(\int_{-T}^{T} Q\left\{\frac{1}{p(t)}\right\} f_1^2(t)p(t)dt, \int_{-T}^{T} f_1^2(t)p(t)dt, \int_{-T}^{T} p(t)dt\right) : p \geqslant 0\right\}.$$

Then note that

(1) Since Q is nonnegative, S is a subset of $\bar{R}^+ \times \bar{R}^+ \times \bar{R}^+$, where $\bar{R}^+ = [0, \infty]$.

(2) Let

$$C = \{(x, y, z) : \exists(x', y', z') \in S \ni x \geqslant x', y = y', z = z'\}.$$

Then C is convex. This follows since the mapping

$$p \to \int_{-T}^{T} Q\left\{\frac{1}{p(t)}\right\} f_1^2(t)p(t)dt$$

is convex while the mappings

$$p \to \int_{-T}^{T} f_1^2(t)p(t)dt, \qquad p \to \int_{-T}^{T} p(t)dt$$

are linear. The convexity of the first mapping was shown in (7).

Let $\bar{C}$ be the closure of C in $\bar{R}^+ \times \bar{R}^+ \times \bar{R}^+$. Let

$$C_{1\eta} = \{(x, \eta) : (x, \eta, 1) \in \bar{C}\},$$
$$C_1 = \{(x, y) : (x, y, 1) \in \bar{C}\}.$$

Let

$$F(x, y) = y^{-1}(1 + 2\gamma x y^{-1}). \tag{15}$$

Note that F is well defined on $C_{1\eta}$ and C_1 since $(x, y, z) \in C$ implies $0 \leqslant y \leqslant Mz$, where $M = \max f_1^2(\cdot)$. Then the minimum value of Problem O is

$$\begin{aligned}\inf\{F(x, y) : (x, y, 1) \in S\} &= \inf\{F(x, y) : (x, y, 1) \in C\} \\ &\geqslant \inf_{C_1} F(x, y).\end{aligned} \tag{16}$$

Similarly, the minimum value of Problem O′ can be related to the infimum of $F(\cdot, \cdot)$ over $\bigcup \{C_{1\eta} : \eta \geqslant \lambda\}$.

We continue the discussion for Problem O; O′ is handled similarly. Since $F(\cdot, \cdot)$ is continuous and C_1 compact, the infimum on the right-hand side of (16) is assumed at $(x(\eta^*), \eta^*)$, say. Of course, $x(\eta^*) < \infty$. In view of the inequality of (16), in order to prove the theorem we need just exhibit a probability density p^* such that

$$\begin{aligned}\int f_1^2 Q\{(p^*)^{-1}\} p^* &= x(\eta^*), \\ \int f_1^2 p^* &= \eta^*.\end{aligned} \tag{17}$$

To establish (17) begin by noting that since the gradient of $F(\cdot, \cdot)$ does not vanish, $(x(\eta^*), \eta^*)$ must lie on the boundary of C_1 and hence $(x(\eta^*), \eta^*, 1)$ lies on the boundary of C. Hence, a supporting hyperplane passes through it, i.e., there exist $\lambda_1, \lambda_2, \lambda_3$ not all 0 such that

$$\lambda_1 x(\eta^*) + \lambda_2 \eta^* + \lambda_3 = \inf\{\lambda_1 x + \lambda_2 y + \lambda_3 z : (x, y, z) \in \bar{C}\}. \tag{18}$$

Note that $(x(\eta^*), \eta^*, 1)$ must also be a boundary point of S since $F(x, y)$ is decreasing in x for fixed y. Thus, there exist $\{p_n\}$ such that $\int p_n \to 1$, $\int f_1^2 p_n \to \eta^*$, $\int f_1^2 Q(p_n^{-1}) p_n \to x(\eta^*)$. In fact, by an application of (3), we can easily show that we may replace p_n by $p_n / \int p_n$ throughout. Thus there exist probability densities q_n such that

$$\int f_1^2 q_n \to \eta^*, \qquad \int f_1^2 Q(q_n^{-1}) q_n \to x(\eta^*). \tag{19}$$

This establishes the identity in (16).

We distinguish several cases in (18).

(i) $\lambda_1 < 0$. This is impossible since $(x', y', z') \in \bar{C}$, $x \geqslant x'$, implies $(x, y', z') \in \bar{C}$ and thus the right-hand side of (18) is $-\infty$.

(ii) $\lambda_1 \geqslant 0$. Then

$$\inf\{\lambda_1 x + \lambda_2 y + \lambda_3 z : (x, y, z) \in \bar{C}\} = \inf\{\lambda_1 x + \lambda_2 y + \lambda_3 z : (x, y, z) \in S\}. \tag{20}$$

(a) $\lambda_1 > 0$. Without loss of generality let $\lambda_1 = 1$ and relabel $\lambda_2 = -\mu$, $\lambda_3 = \mu\tau$. By construction and (4), q_n given by (19) satisfy

$$V(q_n) \to V(q). \tag{21}$$

Therefore, by (5), we must have

$$\int f_1^2 (q_n - q)^2 \to 0. \tag{22}$$

Thus,

$$\int f_1^2 q_n \to \int f_1^2 q \tag{23}$$

and, using (3), we have

$$\int f_1^2 \{Q(q_n^{-1}) q_n - Q(q^{-1}) q\} \to 0. \tag{24}$$

If $\tau \neq 0$, we can combine (21), (23) and (24) to conclude that

$$\int q = 1 \tag{25}$$

and obtain possibility (a) of the theorem. If $\tau = 0$, $q = \{H^{-1}(\mu)\}^{-1}$ on $[-T, T]$. By (22), we get

$$x(\eta^*) = \frac{Q\{H^{-1}(\mu)\}}{H^{-1}(\mu)} \int f_1^2 = \frac{Q\{H^{-1}(\mu)\}}{H^{-1}(\mu)} \int_{[f_1^2 \neq 0]} f_1^2, \qquad \eta^* = \frac{1}{H^{-1}(\mu)} \int f_1^2 = \frac{1}{H^{-1}(\mu)} \int_{[f_1^2 \neq 0]} f_1^2, \tag{26}$$

and

$$\frac{1}{H^{-1}(\mu)} m[f_1^2 \neq 0] \leqslant 1. \tag{27}$$

Let

$$(28)\quad p^*(x) = \frac{1}{H^{-1}(\mu)} \quad \text{if } f_1^2(x) \neq 0,$$
$$= \left(2T - m[f_1^2 \neq 0]\right)^{-1}\left(1 - m[f_1^2 \neq 0]\{H^{-1}(\mu)\}^{-1}\right) \quad \text{otherwise.}$$

Clearly p^* achieves $(x(\eta^*), \eta^*, 1)$ and we have possibility (c) of the theorem.

(b) $\lambda_1 = 0$. We must have $\lambda_2 \geqslant 0$ by arguing as before. It is impossible to have $\lambda_2 = 0$ since $\inf \lambda_3 \int p = 0$ or $-\infty$ and not λ_3. Without loss of generality let $\lambda_2 = 1$. Then

$$(29)\quad \eta^* + \lambda_3 = \inf\{\textstyle\int (f_1^2 + \lambda_3)p : p \geqslant 0\} = -\infty \text{ if } f_1(x) + \lambda_3 < 0 \text{ on a set of positive Lebesgue measure}$$
$$= 0 \quad \text{otherwise.}$$

Since $\eta^* + \lambda_3$ is finite, the first case is impossible. For the second case let $A = \{x : f_1^2(x) = -\lambda_3\}$. Note two subcases:

(b′) $m(A) = 0$,

(b″) $m(A) > 0$.

In both cases if the q_n are defined by (19)

$$(30)\quad \textstyle\int (f_1^2 + \lambda_3) q_n \to 0.$$

It is impossible to have $\lambda_3 = 0$ since then $F\{x(\eta^*), \eta^*\} \geqslant 1/\eta^* = \infty$.

Consider the case (b′). If A_ε is the ε neighborhood of A, we conclude from (30) that

$$(31)\quad \textstyle\int_{A_\varepsilon^c} q_n \to 0$$

and hence that

$$(32)\quad \liminf \textstyle\int f_1^2 q_n^2 \geqslant \liminf \int_{A_\varepsilon} f_1^2 q_n^2$$
$$\geqslant \{m(A_\varepsilon)\}^{-1} \liminf \textstyle\int_{A_\varepsilon} |f_1| q_n$$
$$= \{m(A_\varepsilon)\}^{-1} \inf_{A_\varepsilon} |f_1|.$$

Therefore, since $\lambda_3 \neq 0$,

$$(33)\quad \liminf_n \textstyle\int f_1^2 q_n^2 = \infty.$$

Hence $x(\eta^*) = \infty$ and case (b′) is not possible.

In case (b'') estimate

$$(34)\quad \textstyle\int f_1^2 Q\left(\frac{1}{q_n}\right) q_n \geqslant \left\{\int_{A_\varepsilon} Q\left(\frac{1}{q_n}\right) q_n\right\} \inf_{A_\varepsilon} f_1^2$$
$$\geqslant c_n \inf_{A_\varepsilon} f_1^2 \inf\left\{\textstyle\int_{A_\varepsilon} Q\left(\frac{c_n}{p}\right) p : p \geqslant 0, \int_{A_\varepsilon} p = 1\right\}$$

where $c_n = \int_{A_\varepsilon} q_n$. Apply Lemma 2 to get (34) greater than or equal to

$$c_n \inf_{A_\varepsilon} f_1^2 Q\{m(A_\varepsilon)c_n\}.$$

By (31) $c_n \to 1$. Letting $n \to \infty$ and then $\varepsilon \to 0$, we get

$$(35)\qquad \begin{aligned} x(\eta^*) &= \lim \int f_1^2 Q\left(\frac{1}{q_n}\right) q_n \\ &\geqslant \inf_A f_1^2 Q\{m(A)\} \\ &= |\lambda_3| Q\{m(A)\}. \end{aligned}$$

Let $p^* = I_A / m(A)$. Then

$$(36)\qquad \int f_1^2 p^* = |\lambda_3| = \lim_n \int f_1^2 q_n$$

by (30) and

$$(37)\qquad \int Q\left(\frac{1}{p^*}\right) f_1^2 p^* = Q\{m(A)\}|\lambda_3| = x(\eta^*).$$

Again p^* achieves $(x(\eta^*), \eta^*, 1)$ and we have possibility (*b*) of the theorem. Thus assertions (i) and (ii) of the theorem have been proved.

To prove (iii) begin by noting that minimization of F over C_1 is equivalent to minimization of

$$(38)\qquad W(p) = \frac{\int p\,dt}{\int f_1^2 p\,dt}\left\{1 + 2\gamma \frac{\int Q\left(\frac{1}{p}\right) f_1^2 p\,dt}{\int f_1^2 p\,dt}\right\}$$

over all $p \geqslant 0$. Since the minimizing value of p, p^*, exists it must satisfy

$$(39)\qquad \frac{\partial}{\partial \varepsilon} W\{p^* + \varepsilon(p - p^*)\}|_{\varepsilon=0} \geqslant 0$$

for all $p \geqslant 0$. Equation (39) is of the form

$$(40)\qquad \int \left[f_1^2 \left\{ H\left(\frac{1}{p^*}\right) - \mu \right\} + \mu\tau \right] (p - p^*) \geqslant 0$$

with $q = p^*$ and, after putting $\int p^* = 1$, we obtain

$$(41)\qquad \mu = \frac{1}{2\gamma} + \frac{2\int_{-T}^{T} Q\left(\frac{1}{p^*}\right) f_1^2 p^* dt}{\int_{-T}^{T} f_1^2 p^* dt},$$

$$(42)\qquad \tau = \mu^{-1}\left\{ \frac{\int_{-T}^{T} f_1^2 p^*}{2\gamma} + \int_{-T}^{T} Q\left(\frac{1}{p^*}\right) f_1^2 p^* dt + \int_{-T}^{T} Q'\left(\frac{1}{p^*}\right) f_1^2 dt \right\}.$$

But (41) is just (3.9) and we can show that (42) is implied by (3.8) and (3.9). To see this compute as follows. Using (3.8) and the definition of p^*, we obtain

$$\int f_1^2 H\left(\frac{1}{p^*}\right)p^* = \mu \int f_1^2 p^* - \mu\tau \tag{43}$$
$$= \frac{1}{2\gamma}\int f_1^2 p^* + 2\int Q\left(\frac{1}{p^*}\right) f_1^2 p^* - \mu\tau$$

by (3.9). Substitute in the definition of H and (42) follows.

REMARK. Define (x_0, y_0, z_0) as in the proof above and suppose it corresponds to $q(\cdot, \mu^*, \tau^*)$. If the interval $I = \{y : (x_0, y, z_0) \in S\}$ consists of more than 1 point then $\mu^* \geqslant 0$. To see this note that by definition $y^{-1}(1 + 2\gamma x_0/y)$ is minimized over I by $y = y_0$. Since $x_0 \geqslant 0$, y_0 must be the upper endpoint of I. On the other hand, by construction, y_0 also minimizes $x_0 - \mu^* y + \mu^*\tau^* z_0$ over I and hence $\mu^* \geqslant 0$.

We do not know whether $\tau^* \geqslant 0$ in general. However, as in Example 2, if $\int_{-T}^{T} f_1^{-2}(t)dt = \infty$, no density of the type $q(\cdot, \mu^*, \tau^*)$ exists for $\mu^* > 0$, $\tau^* < 0$.

Acknowledgment. We are indebted to M. Schilling for the remark that the case $\tau < 0$ in Theorem 3.1 can be dispensed with.

REFERENCES

BICKEL, P. J., HERZBERG, A. M and SCHILLING, M. (1977). Robustness of design against autocorrelation in time II: Numerical results for the first order autoregressive process. To be submitted.

BLIGHT, B. N. and OTT, L. (1975). A Bayesian approach to model inadequacy for polynomial regression. *Biometrika* **62** 79–88.

BLOOMFIELD, P. and WATSON, G. S. (1975). The inefficiency of least squares. *Biometrika* **62** 121–128.

BRILLINGER, D. R. (1973). Estimation of the mean of a stationary time series by sampling. *J. Appl. Probability*. **10** 419–431.

COX, D. R. and HINKLEY, D. V. (1968). A note on the efficiency of least-squares estimates. *J. Roy. Statist. Soc. Ser. B* **30** 284–289.

GRENANDER, U. and ROSENBLATT, M. (1957). *Statistical Analysis of Stationary Time Series*. Wiley, New York.

HOEL, P. G. (1958). Efficiency problems in polynomial estimation. *Ann. Math. Statist*. **29** 1134–1145.

HUBER, P. J. (1975). Robustness and designs. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastava, ed.), 287–301. North Holland, Amsterdam.

JEFFREYS, H. (1939). *Theory of Probability*. Clarendon Press, Oxford.

KIEFER, J. (1974). General equivalence theory for optimal designs. *Ann. Statist*. **2** 849–879.

KNOTT, M. (1975). On the minimum efficiency of least squares. *Biometrika* **62** 129–132.

MORRISON, D. F. (1970). The optimal spacing of repeated measurements. *Biometrics* **26** 281–290.

SACKS, J. and YLVISAKER, D. (1966). Designs for regression problems with correlated errors. *Ann. Math. Statist*. **37** 66–89.

SACKS, J. and YLVISAKER, D. (1968). Designs for regression problems with correlated errors; many parameters. *Ann. Math. Statist*. **39** 49–69.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE
QUEEN'S GATE
LONDON SW7 2AZ
ENGLAND