Robustness of Design Against Autocorrelation in Time II: Optimality, Theoretical and Numerical Results for the First-Order Autoregressive Process

P. J. BICKEL, AGNES M. HERZBERG, and M. F. SCHILLING*

The optimum design and its efficiency relative to the uniform design for estimating the mean of a stationary first-order autoregressive process plus an independent error are characterized completely. This is related to the work of Jones (1948). Optimum designs and variances of least squares estimates are calculated numerically for the problem of estimating the slope in a simple linear regression when the errors follow the above structure, for a range of values of the sample size and the parameters of the process. Numerical results in both cases are compared with asymptotic values obtained in Bickel and Herzberg (1979). The asymptotic optimality of the uniform design is borne out.

KEY WORDS: Robustness; Asymptotic design; Autoregressive process.

1. INTRODUCTION

In a recent paper Bickel and Herzberg (1979) (referred to throughout as I) developed a new asymptotic theory for studying the effect of dependence of the observations in the design of experiments for the linear model. Designs that are asymptotically optimal for estimating the parameters under a known dependence structure of the type given below were characterized. The uniform design turned out 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 regardless of the shape of ρ below. The uniform design thus seems to have a claim to robustness against dependence.

In this paper we study the procedures of I numerically for the location and linear regression models when the dependence is first-order autoregressive and give some fixed sample optimality results for the location case.

As in I, suppose that observations can be taken on a variable y at N time points, $-T \leq t_1 \leq \cdots \leq t_N \leq T$, and that an observation at time T can be written

$$Y(t) = \beta_1 f_1(t) + \cdots + \beta_p f_p(t) + \epsilon(t), \qquad (1.1)$$

where the $f_j(t)$ are known functions, the β_j are unknown parameters ($j = 1, \ldots, p$), and $\epsilon(t)$ is a random error with center zero.

If $\epsilon_1(s)$ and $\epsilon_2(s)$ are the errors of two observations taken at times s and t, then we suppose

$$\epsilon_1(s) = \epsilon'(s) + \epsilon_1'',$$

$$\epsilon_2(t) = \epsilon'(t) + \epsilon_2'',$$

where $\epsilon'(\cdot)$ 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 (1961, p. 297), which is discussed further later.

Under this structure, the errors $\epsilon_i(t_i)$ ($i = 1, \ldots, N$), have a joint normal distribution with

$$E\{\epsilon_i(t_i)\} = 0, \qquad (1.2)$$

$$\operatorname{var}\{\epsilon_i(t_i)\} = \sigma^2, \qquad (1.3)$$

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

In I the interest was in situations where the dependence of the observations is not known too precisely and at least initially is assumed to be negligible. For N observations, the correlation function was taken to be given by

$$\rho(t) = \rho_N(t) = \rho_1(Nt), \qquad (1.5)$$

where $\rho_1(t) \to 0$ as $t \to \infty$. Thus $\rho(\cdot)$ depends on N and is close to the correlation function for independent errors if N is large.

The initial assumption that $\rho(\cdot)$ is not known precisely suggested that designs should be studied for which the ordinary least squares estimators, which are optimal when the errors are uncorrelated, perform well. Thus the

* Peter J. Bickel is Professor, Department of Statistics, University of California, Berkeley, CA 94720. Agnes M. Herzberg is Lecturer, Department of Mathematics, Imperial College, London, England. Mark F. Schilling is Assistant Professor, Department of Mathematics, University of Southern California, Los Angeles, CA 90007. Research was supported in part by the National Science Foundation Grant MCS76-10238 and Office of Naval Research Contract No. N00014-75-C-0444/NR 042-036. We are indebted to B. Blight and D. Ylvisaker for bringing the Jones and Chipman et al. references to our attention.

© Journal of the American Statistical Association
December 1981, Volume 76, Number 376
Theory and Methods Section

asymptotic theory of the variance-covariance matrix of the least squares estimates where

$$\sigma^2\hat{\Sigma} = \sigma^2(\mathbf{F}^T\mathbf{F})^{-1}(\mathbf{F}^T\mathbf{U}\mathbf{F})(\mathbf{F}^T\mathbf{F})^{-1}, \quad (1.6)$$

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

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

was investigated.

In Section 2 we consider the location model, $p = 1$, $f_1(t) = 1$. We characterize the design that minimizes (1.6) for fixed ρ and N (Theorem 1) and calculate it explicitly for

$$\rho_1(t) = e^{-\lambda|t|} \quad (\lambda > 0). \quad (1.9)$$

In this model we compare this optimal design to the discretized (asymptotically) optimum uniform design for $N = 10, 20$ and what we argue is a practical range of λ and γ. In Section 3 we consider the regression model, $p = 2$, $f_1(t) = 1$, $f_2(t) = t$ but restrict ourselves to symmetric designs. In Section 3.1 we indicate how the design that asymptotically minimizes the variance of the slope estimate $\hat{\beta}_2$ is computed numerically and give its qualitative shape. The uniform design is still asymptotically optimal here for $\hat{\beta}_1$; for $\hat{\beta}_2$ the uniform design is only optimal as $T \to 0$ (see I). We compare the asymptotic performance of the uniform asymptotically optimal symmetric designs for ρ_1 given by (1.9) with what we call the mimic design. The mimic design has uniform mass equal to $\frac{1}{2}$ supported in each of the intervals $\{-1, -\tau^{1/2}\}$ and $\{\tau^{1/2}, 1\}$, where $\tau^{1/2}$ is determined for the optimal design. In Section 3.2 we sketch the numerical computation of the optimal design for estimating the slope β_2 for fixed N and ρ_1. Numerical comparisons between the exact and approximate and mimic designs are given in Section 3.3 for the correlation function (1.9). Conclusions are given in Section 4 and the Appendix contains the proofs of the theorems in Section 2.

2. LOCATION MODEL

If, in (1.1), $p = 1$ and $f_1(t) \equiv 1$, the asymptotically optimal design on $\{-T, T\}$ is the uniform design, as was shown in I. For fixed N, the optimal design is characterized in the following theorem. Recall that a symmetric design has $t_i = -t_{N-i+1}$ $(i = 1, \ldots, N)$.

Theorem 1. (a) If $\rho_1(t)$ given by (1.5) is convex on $[0, \infty)$ and continuous, an optimal design $\{t_1^*, \ldots, t_N^*\}$ that is symmetric always exists. (b) If, further, $\rho_1(t)$ is strictly convex and differentiable on $(0, \infty)$ with a right derivative at 0, the optimal design is unique and necessarily symmetric. The optimal design satisfies the following conditions uniquely:

1. $-T < t_k^* \le 0$ implies $t_{k-1}^* < t_k^* < t_{k+1}^*$ for all k.
2. There is an $r \ge 1$ such that $t_1^* = \cdots = t_r^* = -T$ and $t_{r+1}^* > -T$.
3. Let $[x]$ represent the greatest integer less than or equal to x. For $r < k \le [\frac{1}{2}(N + 1)]$,

$$\sum_{i=1}^{k-1} \rho_1'\{N(t_k^* - t_i^*)\} = \sum_{i=k+1}^{N} \rho_1'\{N(t_i^* - t_k^*)\}; \quad (2.1)$$

4. $$(r - 1)\rho_1'(0) \ge \sum_{i=r+1}^{N} \rho_1'\{N(t_i^* + T)\}. \quad (2.2)$$

With the aid of Theorem 1, the optimal design is completely specified for $\rho_1(t) = e^{-\lambda|t|}$ and any N, $\lambda > 0$ by the following theorem.

Theorem 2. If $\rho_1(t) = e^{-\lambda|t|}$, the unique optimal design $\{t_1^*, \ldots, t_N^*\}$ satisfies

1. $t_1^* = \cdots = t_r^* = -T$, $t_{r+1}^* = a_N - T$ for $a_N > 0$ for some $r \in \{1, \ldots, [N/2]\}$.
2. $t_{i+1}^* - t_i^* = b_N > 0$ for $i = r + 1, \ldots, [N/2]$ if $r < [N/2]$.

Furthermore, r, a_N, and b_N are specified by the following:

3. If N is even and $\lambda T \le \log(N/(N - 2))/2N$, $r = N/2$. (Hence $a_N = 2T$.)
4. If N is odd and $\lambda T \le \log((N - 1)/(N - 3))/N$, $r = (N - 1)/2$ and $a_N = T$.
5. Otherwise, let $a = N\lambda a_N$, $b = N\lambda b_N$. Then r, a, and b are the unique values satisfying equations

$$2a + (N - 2r - 1)b = 2N\lambda T \quad (2.3)$$

$$e^a = r(e^b - 1) \quad (2.4)$$

$$a \le b. \quad (2.5)$$

The proofs of both theorems are given in the Appendix.

The following corollary, which specifies r for large N, follows readily from Theorem 2.

Corollary. For given λ, T there exists N_λ such that, for $N \ge N_\lambda$, r is constant. The value of r is the unique solution to $(e^{2\lambda T} - 1)^{-1} < r \le e^{2\lambda T}(e^{2\lambda T} - 1)^{-1}$.

From the corollary we see by expanding $e^{2\lambda T}$ that for λT small and sufficiently large N, the optimal design places a total of approximately $1/\lambda T$ points at the ends of the design region. (In fact, it can be shown that for any value of λT, either $r = [1/2\lambda T]$ or $r = [1/2\lambda T] + 1$.) The remaining points are distributed uniformly on $[-T, T]$, with the exception that the first and last spacings are never more, and typically less, than the other spacings.

The optimal designs for location are seen to be independent of the correlation parameter γ. Some examples of optimal designs for $\rho_1(t) = e^{-\lambda|t|}$ are given below:

$\lambda = 1.0$, $N = 10$:

t_i's = ±1, ±.786, ±.562, ±.337, ±.112;

$\lambda = .2, N = 10$:

t_i's = ±1, ±1, ±.751, ±.450, ±.150;

$\lambda = .02, N = 10$:

t_i's = ±1, ±1, ±1, ±1, ±.599.

An intuitive justification for the form of the optimal design as indicated by Theorems 1 and 2 can be given. It is easy to see that the average distance from an observation to the remaining design points is maximized by placing the observation at one of the two endpoints of the design region—and this is desirable. However, the number of sample points that should be placed at an end is limited by the lack of separation incurred between such points, because

1. By the convexity of ρ more is lost by moving one point closer to another one than is gained by moving it away;
2. The number of pairwise differences of zero due to ties at an endpoint increases as the square of the number of points placed there.

The absence of ties or clusters in the interior is also suggested by (1).

2.1 Reduction to $T = 1$

Define the equally spaced design for N observations by $t_i = T\{2(i - 1)/(N - 1) - 1\}$, $(i = 1, \ldots, N)$. Consider generally the optimization problem for correlation functions that depend on t through λt. For given λ and T, let $V_N(\lambda, T) = N \operatorname{var}(\bar{Y})$ for the optimum symmetric design with N observations on the interval $[-T, T]$ for such a correlation function. Similarly, let $U_N(\lambda, T) = N \operatorname{var}(\bar{Y})$ for the equally spaced design. It is evident by matching design $\{t_1, \ldots, t_N\}$ with design $\{t_1T^{-1}, \ldots, t_NT^{-1}\}$ that $V_N(\lambda, T) = V_N(\lambda T, 1)$ and $U_N(\lambda, T) = U_N(\lambda T, 1)$. Moreover, these identities hold for $N = \infty$ as well. Therefore, efficiency comparisons for the two designs may be made for $T = 1$ and selected values of λ and N.

2.2 Discussion of Table 1

Define the efficiency of the equally spaced design by $\{U_N(\lambda, T)/V_N(\lambda, T)\}^{-1}$. Table 1 gives efficiencies for $T = 1$ and for selected values of N, γ, and λ.

The excellent performance of the approximation to the asymptotic result is not surprising since the optimal design is so close in shape to the equally spaced one. It is interesting to note that save for the fixed number of points reserved for ±1 the optimal design is of the same shape as that found by Jones (1948) to be optimal for estimating $\beta_1 + \int_0^1 \epsilon(t)dt$ when $\rho_1(t) = e^{-\lambda|t|}$. Jones's work grew out of design questions in sampling theory. Correspondingly, the results of I should be extendable to the sampling context.

Although our design is optimal only when the nonoptimal least squares estimate is used, the work by Jones, in which the sample mean is optimal for his design, the work by Chipman et al. (1968), and asymptotic calculations indicate that for this ρ_1 the difference is negligible.

We can justify the range of λ and γ we consider as follows. We obtain the same asymptotic efficiencies if, instead of assuming that the interval $[-T, T]$ is fixed and ρ moves, we suppose ρ is fixed, $\rho(t) = \rho_1(t)$, and the possible interval of observation is $[-NT, NT]$. In this context it seems clear that a relevant parameter that essentially does not depend on N is the correlation between successive observations in the equally spaced design, which is approximately $\gamma e^{-2\lambda T}$.

The problem of correlation in observations is discussed by Pearson (1902). Jeffreys (1961) reports that Pearson

> carried out some elaborate experiments to test whether errors of observation could be treated . . . as a combination of a random error with a constant systematic error for each observer. . . . For each type of observation there were three observers who each made about 500 observations. When these observations were taken in groups of 25 to 30 it was found that the means fluctuated not by the amounts that would correspond to the means of 25 to 30 random errors with the general standard error indicated by the whole series but by as much as the means of 2 to 15 observations should.

If we apply our model and formula (3.1) of I we find that the variance of the mean $\bar{X}_N$ of N observations spaced $2T$ apart is approximately $(\sigma^2/N)\{1 + 2\gamma Q(2\lambda T)\} = (\sigma^2/N)\{1 + 2\gamma(e^{2\lambda T} - 1)^{-1}\}$. The mean of N independent observations of course has variance σ^2/N. From the Pearson statement we expect that the ratio of these

Table 1. The Efficiency of the Equally Spaced Design

λ	γ = 1	.5	.2	N
1.0	.9998	.9999	.99995	10
	.9998	.9999	.99996	20
.8	.9989	.9995	.9998	10
	.9994	.9996	.9998	20
.6	.9968	.9980	.9991	10
	.9975	.9984	.9992	20
.4	.9868	.9910	.9954	10
	.9893	.9924	.9960	20
.2	.9461	.9578	.9745	10
	.9536	.9624	.9760	20
.1	.9251	.9367	.9568	10
	.9154	.9255	.9453	20
.08	.9246	.9353	.9547	10
	.9073	.9170	.9369	20
.06	.9291	.9382	.9555	10
	.9024	.9112	.9301	20
.04	.9392	.9462	.9601	10
	.9054	.9125	.9286	20
.02	.9606	.9647	.9731	10
	.9278	.9322	.9427	20

quantities ranges from approximately 12.5 to 2. For $T = 1$ this corresponds to the following ranges for λ and the correlation between successive observations, $\gamma e^{-2\lambda}$.

γ	λ	$\gamma e^{-2\lambda}$
1	.08–.55	.33–.85
.5	.04–.35	.25–.46
.2	.02–.17	.14–.19

3. LINEAR REGRESSION MODEL

3.1 The Case $N = \infty$

Asymptotic designs. If, in (1.1), $p = 2$, $f_1(t) = 1$, $f_2(t) = t$, and one is interested in $N \operatorname{var}(\hat\beta_2)$, it was shown in I under suitable conditions, including the convexity of $\rho(t)$ on $(0, \infty)$, that the asymptotically optimal *symmetric* design q is a solution to the equations

$$\int_{-T}^{T} q(t, \mu^*, \tau^*)dt = 1, \tag{3.1}$$

$$\frac{2\int_{-T}^{T} Q\left(\frac{1}{q}\right) t^2 q dt}{\int_{-T}^{T} t^2 q dt} = \mu^* - \frac{1}{2\gamma}, \tag{3.2}$$

where, for $\mu > 0$, $0 \le \tau < T^{1/2}$,

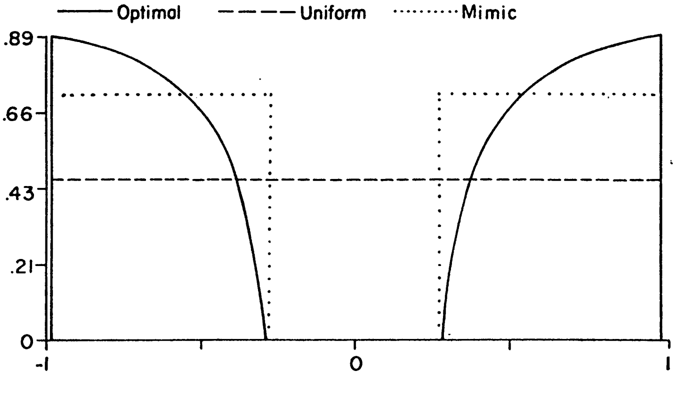
$$q(t, \mu, \tau) = \begin{cases} 0 & (|t| \le \tau^{1/2}), \\ [H^{-1}\{\mu(1 - \tau t^{-2}\}]^{-1} & (\tau^{1/2} \le |t| \le T), \\ 0 & (|t| > T), \end{cases} \tag{3.3}$$

$$Q(t) = \sum_{j=1}^{\infty} \rho_1(jt)$$
$$H(t) = Q(t) - tQ'(t). \tag{3.4}$$

Such a density is given in Figure 1.

Unfortunately, we were unable to show that the solution to these equations is unique. However, we did obtain solutions that yielded designs with asymptotic variances below those of any competitors we considered. We conjecture that these solutions are the unique asymptotically optimal designs. These solutions are the source of Table 2.

Figure 1. Optimal, Uniform, and Mimic Designs for $T = 1$, $\gamma = .5$, $\lambda = .483$, and $\tau^{1/2} = .2985$



In addition, we consider two suboptimal designs, the uniform design and the mimic design, both defined in Section 1. Although the mimic design can be defined for arbitrary $|\tau| \le T^2$, we only give numerical results for the interesting case when it indeed mimics the optimal design, that is, $\tau = \tau^*(\lambda)$. Figure 1 shows comparable optimal, uniform, and mimic designs for $\tau^{1/2} = .2985$, $\gamma = .5$, and $\lambda = .483$.

Reduction to $T = 1$. Let $\tilde V(\lambda, T)$ be the asymptotic variance of $N^{1/2}\hat\beta_2$ for the optimum symmetric design. Arguing as in the location case, it follows readily that $\tilde V(\lambda, T) = T^{-2}\tilde V(\lambda T, 1)$. The same relationship holds for both the uniform and mimic designs. Moreover, if we assume that (3.1) and (3.2) have a unique solution for each λ, we readily get (indicating now the dependence of μ^* and τ^* on T) $\mu^*(\lambda T, 1) = \mu^*(\lambda, T)$ and $\tau^*(\lambda T, 1) = T^{-2}\tau^*(\lambda, T)$. Thus all our efficiency comparisons can again be carried out for $T = 1$ and selected λ.

Computations. Simple expressions exist for the variances of the uniform and mimic designs. From Formula (3.1) of I, the asymptotic variance of $N^{1/2}\hat\beta_2$ for a design whose design measure has distribution function F is

$$V(a) = (\sigma^2/D)(1 + 2\gamma C/D) \tag{3.5}$$

where

$$a(t) = F^{-1}(t),$$
$$C = \int_0^1 Q\{a'(t)\}a^2(t)dt,$$
$$D = \int_0^1 a^2(t)dt.$$

For ρ_1 given by (1.9)

$$Q(t) = (e^t - 1)^{-1}. \tag{3.6}$$

Using $a(t) = 2t - T$, we obtain for the uniform design with $\rho_1(t) = e^{-\lambda|t|}$ and $T = 1$,

$$V(2t - 1) = 3\{1 + 2\gamma(e^{2\lambda} - 1)^{-1}\}\sigma^2, \tag{3.7}$$

and for the mimic design,

$$a(t) = \begin{cases} 2(1 - \tau^{1/2})t - 1 & (0 \le t \le \frac{1}{2}), \\ 2(1 - \tau^{1/2})t + (2\tau^{1/2} - 1) & (\frac{1}{2} < t \le 1), \end{cases}$$

yielding

$$V(a) = \frac{3\sigma^2}{(1 + \tau^{1/2} + \tau)} \times (1 + 2\gamma[\exp\{2\lambda(1 - \tau^{1/2})\} - 1]^{-1}). \tag{3.8}$$

Details of the computation of the conjectured asymptotically optimal design from (3.1) and (3.2) are available from the authors upon request.

3.2 The Case $N < \infty$

Designs and computational methods. If $t_1 \le \cdots \le t_N$ is a design and $\rho_1(t) = e^{-\lambda|t|}$, the variance of $N^{1/2}\hat\beta_2$ is given by

$$V(t_1, \ldots, t_N; \lambda) = \frac{N\sigma^2}{\sum_{i=1}^{N} t_i^2} \times \left\{ 1 + \frac{2\gamma \sum_{i<j} t_i t_j e^{-\lambda N(t_j - t_i)}}{\sum_{i=1}^{N} t_i^2} \right\}.$$

Clearly, $V(t_1, \ldots, t_N, \lambda) = T^{-2} V(t_1 T^{-1}, \ldots, t_N T^{-1}, \lambda T)$, and we can again restrict to $T = 1$. The design points for the various designs are

1. Uniform design:

$$t_i = -1 + 2(i-1)/(N-1) \quad (i = 1, \ldots, N).$$

2. Mimic design:

$$t_i = \begin{cases} -1 + \frac{2}{N}(1 - \tau^{1/2})(i-1) & (i = 1, \ldots, \frac{1}{2}N), \\ -t_{N+1-i} & (i = \frac{1}{2}N + 1, \ldots, N). \end{cases}$$

3. Approximate optimal design: Let G be the distribution function of the asymptotically optimal design given by (3.1) through (3.3). The approximate optimal design has $t_i = G^{-1}((i-1)/(N-1))$ $(i = 1, \ldots, N)$. The approximate optimal design is clearly symmetric.

4. Exact optimal design (symmetric): We seek design points $-1 \le t_1 \le \cdots \le t_N \le 1$ such that $t_i = -t_{N-i+1}$ and $V(t_1, \ldots, t_N; \lambda)$ is a minimum. We arrive at a system of equations and inequalities similar to (2.1) and (2.2) by differentiating with respect to t_k, $k = 1, \ldots, M$ and requiring the derivatives to be zero in the interior and nonnegative at the boundaries and specifying $t_1 = \cdots = t_r = -T$.

Unfortunately, we cannot show that the solution is unique or that every solution yields the optimal design. Starting from the uniform case, we solved the equations iteratively and again found that in every case the value of $V(t_1, \ldots, t_N, \lambda)$ obtained from the algorithm, which we call $\check{V}_N(\lambda, T)$, is smaller than the variances of all the competitors.

The optimal designs found consisted of two components: (a) several values are placed at the endpoints of the design region; and (b) the values in the interior correspond approximately to the asymptotically optimal design measure. Note that this is precisely what was found for the location designs as well.

An example of a design for regression is the following, optimal for $\lambda = .4$, $\gamma = 1$, $N = 20$: ± 1, ± 1, $\pm .926$, $\pm .839$, $\pm .751$, $\pm .663$, $\pm .574$, $\pm .482$, $\pm .389$, $\pm .291$.

3.3 Tables

Table 2 gives values of $\tau^{1/2}$ and the corresponding optimal, uniform, and mimic asymptotic variances of $N^{1/2}\hat{\beta}_2$, with the percentage due to the dependence component

$$1 - \left\{ V(a) \int_0^1 a^2(t)dt \right\}^{-1}$$

for different values of γ and λ. The dependence component increases as λ decreases. Table 2 also gives the efficiencies of the uniform and mimic designs to the optimal design in terms of the ratio of asymptotic variances.

Table 2. Asymptotic Variances and Efficiencies of $N^{1/2}\hat{\beta}_2$ for Various Designs When $\rho_1(t) = e^{-\lambda|t|}$, $T = 1$, and $\sigma = 1$

λ	$\tau^{1/2}$	Optimal Variance	% Dependence Component	Uniform Variance	Mimic Variance	Optimal/Uniform	Optimal/Mimic
				$\gamma = 1$			
1.0	.2840	3.5334	43.48	3.9391	3.5779	.8970	.9876
.8	.2366	4.2208	50.03	4.5178	4.2614	.9343	.9905
.6	.1843	5.3934	58.30	5.5861	5.4269	.9655	.9938
.4	.1270	7.7969	68.85	7.8967	7.8193	.9874	.9971
.2	.0653	15.1721	82.42	15.2016	15.1807	.9981	.9994
				$\gamma = .5$			
1.0	.4260	2.6700	37.14	3.4695	2.7335	.7695	.9768
.8	.3837	3.0575	42.71	3.7589	3.1256	.8134	.9782
.6	.3331	3.6996	50.06	4.2930	3.7717	.8618	.9809
.4	.2710	4.9754	60.18	5.4479	5.0497	.9133	.9853
.2	.1890	8.7730	74.96	9.1008	8.8430	.9640	.9921
				$\gamma = .2$			
1.0	.5785	1.9804	28.16	3.1878	2.0421	.6212	.9698
.8	.5420	2.1697	32.25	3.3036	2.2390	.6568	.9691
.6	.4949	2.4731	37.90	3.5172	2.5520	.7031	.9691
.4	.4301	3.0525	46.38	3.9793	3.1439	.7671	.9709
.2	.3285	4.6927	61.08	5.4399	4.7997	.8627	.9777

Table 3. Variances and Efficiencies of $N^{1/2}\hat{\beta}_2$ for Various Designs and Values of N When $\rho_1(t) = e^{-\lambda|t|}$, γ = .5, T = 1, and σ = 1

λ	$\tau^{1/2}$	N	Optimal	Approximate Optimal	Mimic	Uniform	Optimal / Appr. Opt.	Optimal / Mimic	Optimal / Uniform	2r
1.0	.4260	10	2.0510	2.0980	2.1283	2.6530	.9776	.9637	.7731	2
		20	2.3325	2.3629	2.4086	3.0266	.9872	.9685	.7708	2
		40		2.5115		3.2385				
		300		2.6487		3.4376				
.8	.3837	10	2.2076	2.2921	2.3261	2.7753	.9631	.9490	.7954	2
		20	2.5838	2.6401	2.6915	3.2192	.9787	.9600	.8026	2
		40		2.8415		3.4758				
		300		3.0280		3.7195				
.6	.3331	10	2.4064	2.5746	2.6137	2.9823	.9347	.9207	.8069	4
		20	2.9527	3.0713	3.1312	3.5618	.9614	.9430	.8290	4
		40		3.3712		3.9064				
		300		3.6544		4.2389				
.4	.2710	10	2.6835	3.0121	3.0552	3.3368	.8910	.8784	.8043	4
		20	3.5670	3.8318	3.9069	4.2345	.9309	.9130	.8424	4
		40	4.1904	4.3663	4.4418	4.7981	.9597	.9434	.8733	4
		300		4.8912		5.3561				
.2	.1890	10	2.9118	3.5994	3.6306	3.8190	.8089	.8020	.7624	8
		20	4.6000	5.4581	5.5489	5.7609	.8428	.8290	.7985	8
		40	6.2440	6.9032	7.0141	7.2444	.9045	.8902	.8619	8
		300		8.5055		8.8323				

As can be seen from the table, the mimic design does very well. Here $\tau^{1/2}$ is optimal for the given λ. As λ increases, $\tau^{1/2}$ tends asymptotically to one. This is obvious on intuitive grounds, for the larger the value of λ, the less dependent the observations; in the independence case $\lambda = \infty$, one-half of the observations are taken at -1, and one-half the observations at $+1$.

Table 3 gives for $\gamma = .5$ and particular values of λ the variances of $N^{1/2}\hat{\beta}$ for the exact design (the optimal design for N points), the approximate optimal design, the mimic design, and uniform (equally spaced) designs and their efficiencies for several values of N. We also show the number $2r$ of points placed by the exact design at ± 1. One can see from the table how quickly the asymptotics work.

Table 3 also shows that on the whole the ordering of the variances of the designs is as one would expect. However, toward the lower end of the range of plausible values of λ, the uniform design begins to perform comparatively poorly again and the optimal design places a large number of points at ± 1.

Interestingly, the number of points at ± 1 is constant in N for each λ in Table 3; in calculations for $\gamma = 1.0$ and .2 and the same values of λ and N, this was also the case. We conjecture that, as in the location case, the number of such points is constant for all N sufficiently large.

4. CONCLUSIONS

The asymptotics of I suggested that the equally spaced design should be nearly optimal for estimating both intercept and slope in a linear regression when the errors have a convex correlation function. Theorem 2 shows that for $\rho_1(t) = e^{-\lambda|t|}$ and location the exact design is nearly uniform for relatively small N. Our numerical results for the first-order autoregressive process seem to bear out the asymptotics for estimating both parameters for reasonable sample sizes and values of the parameters of the process. They also suggest that if the order of magnitude of the correlation is known, linearity of the response is secure, and estimation of the slope is desired, the mimic designs for $\tau > 0$ may be worth pursuing.

APPENDIX

Proof of Theorem 1. (a) Let $V(\mathbf{t}) = V(t_1, \ldots, t_N) = N \operatorname{var}(\bar{Y})$ when the design points are $t_1 \le \cdots \le t_N$ and $\bar{Y}$ is the mean of the observations. It is easy to see that

$$V(\mathbf{t}) = \sigma^2\left[1 + 2\gamma N^{-1} \sum_{j=2}^{N} \sum_{i=1}^{j-1} \rho_1\{N(t_j - t_i)\}\right], \quad \text{(A.1)}$$

where σ^2 is the experimental error variance. Our objective is to minimize $V(\mathbf{t})$ on $S = [\{t_1, \ldots, t_N\}\colon -T \le t_1 \le \cdots \le t_N \le T]$. Since $V(\mathbf{t})$ is continuous and S is compact, a minimum point clearly exists.

Suppose $\{t_1', \ldots t_N'\}$ is such a point. Then, by the symmetry of $V(\mathbf{t})$, so is $\{-t_N', \ldots, -t_1'\}$. Since $\rho_1(\cdot)$ is convex on $[0, \infty)$, $V(\mathbf{t})$ is convex on S and hence the symmetric design $\{t_1^*, \ldots, t_N^*\}$, where $t_i^* = 1/2(t_i' - t_{N-i+1}')$ $(i = 1, \ldots, N)$ is also a minimum point for $V(\mathbf{t})$. Part (a) follows.

To prove (b), note first that $V(\mathbf{t})$ is strictly convex and unicity and symmetry of the optimal design follow. By convexity of $V(\mathbf{t})$ and differentiability of $\rho_1(\cdot)$ on $[0, \infty)$, a necessary and sufficient condition for $\{t_1^*, \ldots, t_N^*\}$

$\in S$ to be a minimum point is that

$$\frac{\partial}{\partial \eta} V\{\mathbf{t}^* + \eta(\mathbf{t} - \mathbf{t}^*)\}]_{\lambda=0} \geq 0 \tag{A.2}$$

for $\mathbf{t}^* = (t_1^*, \ldots, t_N^*)$ and every $\mathbf{t} \in S$. The left-hand side of expression (A.2) easily reduces to

$$2\gamma\sigma^2 \sum_{j=1}^{N} \sum_{i=1}^{j-1} \{(t_j - t_i) - (t_j^* - t_i^*)\}\rho_1'\{N(t_j^* - t_i^*)\}$$
$$= 2\gamma\sigma^2 \sum_{j=1}^{N} (t_j - t_j^*) \times \left[\sum_{i=1}^{j-1} \rho_1'\{N(t_j^* - t_i^*)\} - \sum_{i=j+1}^{N} \rho_1'\{N(t_i^* - t_j^*)\}\right], \tag{A.3}$$

where sums with overlapping indices are interpreted as 0.

1. We show that $-T < t_{j_0}^* < T$ implies $t_{j_0-1}^* < t_{j_0}^* < t_{j_0+1}^*$; that is, there are no ties in the interior of the design region. For a symmetric design, this is just (1). Suppose there are ties in the interior; that is, there exists j_0 and $k \geq 1$ such that

$$-T < t_{j_0}^* = \cdots = t_{j_0+k}^* < T.$$

Taking $t_j = t_j^*$ $(j \neq j_0)$, $t_{j_0} < t_{j_0}^*$, we obtain from (A.3)

$$\sum_{i=1}^{j_0-1} \rho_1'\{N(t_{j_0}^* - t_i^*)\} - \sum_{i=j_0+k+1}^{N} \rho_1'\{N(t_i^* - t_{j_0}^*)\} - k\rho_1'(0) \leq 0. \tag{A.4}$$

Similarly, taking $t_j = t_j^*$ $(j \neq j_0 + k)$, $t_{j_0+k} > t_{j_0+k}^*$, we get

$$\sum_{i=1}^{j_0-1} \rho_1'\{N(t_{j_0}^* - t_i^*)\} - \sum_{i=j_0+k+1}^{N} \rho_1'\{N(t_i^* - t_{j_0}^*)\} + k\rho_1'(0) \geq 0. \tag{A.5}$$

From (A.4) and (A.5) we have that $-k\rho_1'(0) \leq k\rho_1'(0)$. Since $\rho_1(\cdot)$ is convex and decreasing on $[0, \infty)$, $\rho_1'(0) < 0$ and we have a contradiction.

2. Suppose $t_N^* = -t_1^* = aT$. If $0 < a < 1$, it is evident from (A.1) that the design with each $t_i = t_i^*/a$ produces a smaller variance since ρ_1 is decreasing, a contradiction. Taking $a = 0$ gives a design clearly inferior to any other design.

3–4. From expression (A.3) we see that since $t_{k-1}^* < t_k^* < t_{k+1}^*$, we can by selecting $t_k \gtrless t_k^*$ force (2.1). The necessity of (4) is similarly clear.

On the other hand, since $t_r^* = -T$, any point $(t_1, \ldots, t_N) \in S$ must have $t_r \geq t_r^* = -T$ and similarly $t_{N-r+1} \leq t_{N-r+1}^* = T$. The sufficiency of (3) and (4) and uniqueness of r follow from (A.2).

Proof of Theorem 2. Part (1) follows directly from Theorem 1. For (2) through (5) we have from Theorem 1 the following necessary and sufficient conditions for the optimal design

$$\sum_{i=k+1}^{N} e^{-N\lambda(t_i^* - t_k^*)} = \sum_{i=1}^{k-1} e^{-N\lambda(t_k^* - t_i^*)} \qquad (r < k \leq [\tfrac{1}{2}(N + 1)]), \tag{A.6}$$

$$re^{-2N\lambda T} + I_{[r<N/2]} \sum_{i=r+1}^{N-r} e^{-N\lambda(t_i^* - t_r^*)} \geq r - 1, \tag{A.7}$$

where $I_{[\cdot]}$ represents the indicator function. We argue cases (3), (4), and (5) successively. If the conditions of (3) hold, then for the design satisfying (1) with $r = N/2$, (A.7) is easily seen to be satisfied and (A.6) is vacuous. As for (4), the design satisfying (1) with $r = (N - 1)/2$ and $a_N = T$ gives for (A.7) $e^{-N\lambda T} + ((N - 1)/2)\, e^{-2N\lambda T} \geq (N - 3)/2$, which easily reduces to the condition of (4). Equation (A.6) follows from symmetry.

Now suppose we have a design satisfying (1) and (2) for some $r < [N/2]$. The condition $t_N^* - t_1^* = \sum_{i=1}^{N-1} (t_{i+1}^* - t_i^*) = 2T$ produces (2.3). Equation (A.6) for this design becomes

$$\sum_{j=1}^{N-k-r} e^{-jb} + re^{-\{a+(N-k-r)b\}} = re^{-\{a+(k-r-1)b\}} + \sum_{j=1}^{k-r-1} e^{-jb} \qquad \text{for } r < k \leq \left[\frac{N}{2}\right],$$

which reduces after some algebra to (2.4).

Equation (A.7) also contains a geometric series for such a design, which allows (2.12) to be expressed as $re^{-2N\lambda T} + e^{-a}(1 - e^{-(N-2r)b})(1 - e^{-b})^{-1} \geq r - 1$. Upon appropriate substitutions of (2.3) and (2.4), this equation can be reduced to just (2.5).

To complete the proof it is necessary to show that for any N and λ not covered by cases (3) and (4), a solution to (2.3), (2.4), and (2.5) exists for some $r < [N/2]$. Write

$$b = h_1(a) = (2N\lambda T - 2a)(N - 2r - 1)^{-1}$$
$$b = h_2(a) = \log(e^a r^{-1} + 1)$$

as equivalent versions of (2.3) and (2.4). Define

$$\ell = \ell(r) = \begin{cases} \ln(r(r - 1)) & (r > 1), \\ +\infty & (r = 1). \end{cases}$$

Since $h_2(a)$ is continuous, increasing, and satisfies (2.5) on $(0, \ell(r)]$, and $h_1(a)$ is continuous and decreasing, a solution exists for given r if and only if $h_1(0) > h_2(0)$ and $h_2(\ell) = \ell \geq h_1(\ell)$. Combining these two conditions gives

$$(N - 2r - 1)\ell(r + 1) < 2N\lambda T \leq (N - 2r + 1)\ell(r). \tag{A.8}$$

Observe that the bounds in (A.8) partition the possible range for $2N\lambda T$ in case (5) as r varies from 1 to $[N/2] - 1$. Hence a solution to (2.3)–(2.5) exists.

[*Received February 1980. Revised April 1981.*]

REFERENCES

BICKEL, P.J., and HERZBERG, A.M. (1979), "Robustness of Design Against Autocorrelation in Time I: Asymptotic Theory, Optimality for Location and Linear Regression," *Annals of Statistics*, 7, 77–95.

CHIPMAN, J., KADIYALA, K.R., MADANSKY, A., and PRATT, J.W. (1968), "Efficiency of the Sample Mean When Residuals Follow a First Order Stationary Markoff Process," *Journal of the American Statistical Association*, 63, 1237–1246.

JEFFREYS, H. (1961), *Theory of Probability* (3rd ed.), Oxford: Clarendon Press.

JONES, A.E. (1948), "Systematic Sampling of Continuous Parameter Populations," *Biometrika*, 35, 283–290.

MORRISON, D.F. (1970), "The Optimal Spacing of Repeated Measurements," *Biometrics*, 26, 281–290.

PEARSON, K. (1902), "C. On the Mathematical Theory of Errors of Judgement, With Special Reference to the Personal Equation," *Philosophical Transactions of the Royal Society of London*, Series A 198, 235–299.