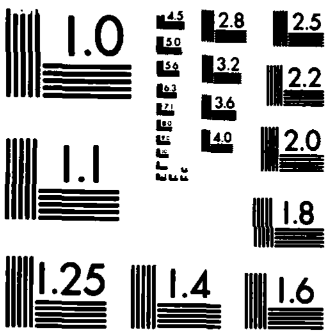
AD-A160 970 ASYMPTOTIC INFERENCE FROM SEQUENTIAL DESIGN IN A NONLINEAR SITUATION(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER C F WU AUG 85 MRC-TSR-2853 1/1

UNCLASSIFIED DAAG29-80-C-0041 F/G 12/1 NL

END

FILMED

DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3

MRC Technical Summary Report #2853

ASYMPTOTIC INFERENCE FROM SEQUENTIAL DESIGN IN A NONLINEAR SITUATION

C. F. J. Wu

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

August 1985

(Received July 30, 1985)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

ASYMPTOTIC INFERENCE FROM SEQUENTIAL DESIGN IN A NONLINEAR SITUATION

C. F. J. Wu

Technical Summary Report #2853
August 1985

ABSTRACT

~~It is shown~~ that confidence regions constructed by the repeated-sampling principle are asymptotically valid for sequential designs in general linear models and nonlinear parameters. The related questions of consistency of parameter estimators and convergence of sequential design to an optimal design are answered positively. An empirical finding of Ford and Silvey (1980) is given a theoretical justification.

AMS (MOS) Subject Classifications: 62L05, 62M10

Key Words: Martingale, Nonlinear design, Optimal design, Repeated-sampling principle, Sequential design

Work Unit Number 4 - Statistics and Probability

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

For estimation of parameters in nonlinear models or nonlinear parameters in linear models, sequential design of experiment is often used to best utilize the information. It results in saving the number of runs. After the termination of the experiment with a fixed sample size, inference (such as hypothesis testing or confidence interval) about the parameter is made. The classical repeated-sampling principal of inference can not be applied because it relies on the repetition of the same design while in the sequential setting it is not repeatable. By using the martingale as a technical tool, it is shown that, at least for large samples, such inference is still justified. The companion questions of consistency of parameter estimators and convergence of sequential design to an optimal design are also answered. Keywords: Statistics; probability,

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

ASYMPTOTIC INFERENCE FROM SEQUENTIAL DESIGN IN A NONLINEAR SITUATION

C. F. J. Wu

1. INTRODUCTION

A major difficulty in designing a nonlinear experiment is that the performance of design depends on the unknown parameters. To utilize the information fully, the experiment has to be conducted sequentially. The choice of the next design point is determined by the estimate of the unknown parameters based on the observations made to date; see, for example, Box & Hunter (1965). Since the data thus generated are dependent and the design points are not repeatable, it is not clear whether the repeated sampling principle of inference can be applied here. Similar inferential questions also arise in other contexts (Cox, 1982; Siegmund, 1980).

Ford & Silvey (1980) studied this question in a special example. Their simulation study indicates that standard confidence intervals, constructed by pretending that the design points were predetermined, perform very well. In §2 we provide a theoretical justification of this empirical finding. In §3 we consider the general problem of sequential design and inference when the parameter of interest is a nonlinear smooth function of the linear parameters in a general linear model. Three issues to be studied are:

(A) consistency of the parameter estimator;

(B) asymptotic validity of the standard procedures for confidence region;

(C) convergence of the sequential design, properly normalized, to an optimal design.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

Details are in §3. The answer to them is yes under quite weak conditions. Crucial to our investigation is a martingale structure underlying the problem. Issue (ii) in small samples was studied in an unpublished manuscript by Ford, Titterington & Wu.

2. A SIMPLE EXAMPLE

Ford & Silvey (1980) considered the design problem for estimating the nonlinear function $g(\theta) = -\theta_1/(2\theta_2)$ in the linear model

$$y = \theta_1 u + \theta_2 u^2 + e = \theta^T v + e, \; v = (u, u^2)^T$$

where y is an observed response corresponding to a control variable at level u, $u \in [-1,1]$, and e is an independent $N(0,1)$ error. Take the first two observations at $u = \pm 1$. For $r > 2$, let $\hat{\theta}_r = (\hat{\theta}_{r1}, \hat{\theta}_{r2})$ be the maximum likelihood estimator of θ based on the first r observations, $\hat{g}_r = g(\hat{\theta}_r)$, and $J_r = v_1 v_1^T + \ldots + v_r v_r^T$ be the corresponding information matrix, $v_r = (u_r, u_r^2)^T$. By maximization of the Gâteaux derivative at $\hat{\theta}_r$ of the Fisher information of $\hat{g}_r$, the next design point u_{r+1} is chosen, from $[-1,1]$, to maximize

$$d_r(u) = (v^T J_r^{-1} c_{\hat{g}_r})^2, \; c_g = (1, 2g)^T \quad .$$

It turns out that u_{r+1} must be 1 or -1.

Suppose that, among the first n observations, s_n are taken at $u = -1$ and $n - s_n$ at $u = 1$ with their means denoted by $\bar{y}_-$ and $\bar{y}_+$. Note that s_n is random. Ford & Silvey (1980) showed that

$$\bar{y}_+ \to \phi_1 = \theta_2 + \theta_1 \quad , \quad \bar{y}_- \to \phi_2 = \theta_2 - \theta_1 \quad , \qquad (2.1)$$

with probability 1 and

$$s_n/n \to \eta_\theta^*(-1) = 1 - \eta_\theta^*(1) = |\theta_2+\theta_1|/(|\theta_2+\theta_1| + |\theta_2-\theta_1|) \quad , \qquad (2.2)$$

which is the probability placed at -1 by the optimal continuous design η_θ^* that minimizes $c_g^T M^-(\eta)c_g$ over η, where $M^-(\eta)$ is a g-inverse of the normalized information matrix $M(\eta) = E(vv^T)$ with η being a probability measure on u over $[-1,1]$. Note that $M(\eta)$ is singular if and only if $|g(\theta)| = \frac{1}{2}$. For $g(\theta) = -\frac{1}{2}$, $\eta_\theta^*(-1) = 1$; for $g(\theta) = \frac{1}{2}$, $\eta_\theta^*(1) = 1$. The strong consistency of the maximum likelihood estimator

$$\hat{\theta}_n = (\hat{\theta}_{n1},\hat{\theta}_{n2}) = \frac{1}{2}\left(\bar{y}_+ - \bar{y}_-, \bar{y}_+ + \bar{y}_-\right) \to \theta \qquad (2.3)$$

follows from (2.1).

Can confidence intervals for g be constructed in the usual manner? The answer is not so obvious since the observations are dependent as a result of the sequential generation of the design points. Repeated sampling of the sequential design results in different choices of the design points $\{u_r\}$, which makes the distribution calculus quite intractable. If the pretence were made that the design was chosen a priori, standard theory would give (Ford & Silvey, 1980, (5.2))

$$\hat{g}_n \pm \hat{\theta}_{n2}^{-1}(c_{\hat{g}_n}^T J_n^{-1} c_{\hat{g}_n})^{1/2} \qquad (2.4)$$

as an approximate 95% confidence interval for g. An alternative to (2.4) is to replace J_n by $nM(\eta^*_{\hat{\theta}_n})$ since, from (2.2), $J_n/n \to M(\eta_\theta^*)$. The two versions are asymptotically equivalent. The latter was shown to perform remarkably well in the empirical study of Ford & Silvey (1980). The empirical percentage coverages of the true parameter are quite close to 95%. A theoretical justification for (2.4) is now in order.

From (2.3) and $J_n/n \to M(\eta_\theta^*)$, the asymptotic validity of (2.4) can be established via the asymptotic normality of the normalized statistic

$$\frac{\sqrt{n}(\hat{g}_n - g)}{(2\theta_2)^{-1}(c_g^T M^-(\eta_\theta^*)c_g)^{1/2}} \to N(0,1) \quad . \tag{2.5}$$

We shall give the proof separately for singular and nonsingular $M(\eta_\theta^*)$.

First consider $g(\theta) = -\frac{1}{2}$. The treatment of $g(\theta) = \frac{1}{2}$ is similar. Since $\phi_1 = 2\theta_2$ and $\phi_2 = 0$, the numerator of (2.5) equals $\sqrt{n}(\bar{y}_- - \phi_2)/(\bar{y}_+ + \bar{y}_-)$, which can be approximated by $\sqrt{n}\ \bar{y}_-(2\theta_2)^{-1}$ via (2.1). Since $\eta_\theta^*(-1) = 1$, the denominator of (2.5) is $(2\theta_2)^{-1}$ and (2.5) can be approximated by $\sqrt{n}\ \bar{y}_-$, whose asymptotic normality follows from the central limit theorem on $\sqrt{s_n}\ \bar{y}_-$ and $s_n/n \to \eta_\theta^*(-1) = 1$. Here we use the fact that, given s_n, the observations taken at $u = -1$ are independent and identically distributed.

For $|g(\theta)| \neq \frac{1}{2}$, a more general result will be proved. Note that

$$\hat{g}_n = \frac{1}{2}\left(\bar{y}_- - \bar{y}_+\right)/\left(\bar{y}_- + \bar{y}_+\right) = A(\bar{y}_+, \bar{y}_-)$$

is a smooth function of $\bar{y}_-$ and $\bar{y}_+$. Similarly, $g = A(\phi_1, \phi_2)$, where $\phi_1 = Ey_+$, $\phi_2 = Ey_-$. From the smoothness of A and (2.1), the asymptotic distribution of $\sqrt{n}(\hat{g}_n - g)$ is given by that of its first order approximation

$$A_1(\phi)\sqrt{n}(\bar{y}_+ - \phi_1) + A_2(\phi)\sqrt{n}(\bar{y}_- - \phi_2) \quad , \tag{2.6}$$

where $A_1(\phi)$ and $A_2(\phi)$ are the partial derivatives of A at $\phi =$

(ϕ_1,ϕ_2) with respect to ϕ_1 and ϕ_2. The denominator of (2.5) equals

$$A_1^2(\phi)/\eta_\theta^*(1) + A_2^2(\phi)/\eta_\theta^*(-1) \quad . \tag{2.7}$$

Therefore, (2.5) would follow from the asymptotic normality of the ratio of (2.6) and (2.7), which is an easy consequence of

$$\sqrt{n}\ a_1(\bar{y}_+ - \phi_1) + \sqrt{n}\ a_2(\bar{y}_- - \phi_2) \to N\left(0, \frac{a_1^2}{\eta_\theta^*(1)} + \frac{a_2^2}{\eta_\theta^*(-1)}\right) \tag{2.8}$$

in distribution for any a_1 and a_2, whose proof is given in the Appendix.

It is obvious from the arguments that the normality assumption on e in the linear model is not essential.

3. GENERAL PROBLEM

The above example is simple and special in that the observations are always taken at $u = \pm 1$. Similar results will be obtained in this section for a more general problem under additional assumptions.

We consider the general linear model

$$y = x^T\theta + \varepsilon$$

where θ is a $p \times 1$ vector, and the design variable x can be chosen anywhere within a bounded design region X. Assumptions on ε are given in (3.3). The $q \times 1$ vector parameter of interest is $\psi = g(\theta)$, which is a nonlinear smooth function of θ. Let $\hat{\theta}_n$ be the least squares estimator of θ based on the first n observations (y_i,x_i). The variance-covariance matrix of $\hat{\psi}_n = g(\hat{\theta}_n)$ is approximated by $\sigma^2 g'(\theta)^T M_n^{-1} g'(\theta)$, where $M_n = x_1x_1^T + \ldots + x_nx_n^T$ and $g'(\theta)$

is the derivative of g. The next design point x_{n+1} is chosen from $x \in X$ to minimize

$$\Phi(M_n + xx^T, \hat{\theta}_n) = \phi(g'(\hat{\theta}_n)^T (M_n + xx^T)^{-1} g'(\hat{\theta}_n)) \quad , \qquad (3.1)$$

where the "optimality criterion" ϕ is a scalar function. For the example in §2, $q = 1$ and ϕ is the identity map. Another choice of x_{n+1} is to minimize the Fréchet derivative of Φ at M_n and $\hat{\theta}_n$ in the direction xx^T (Silvey, 1980), that is,

$$\lim_{\lambda \to 0^+} \lambda^{-1} [\Phi\{(1-\lambda)M_n + \lambda xx^T, \hat{\theta}_n\} - \Phi(M_n, \hat{\theta}_n)] \quad . \qquad (3.2)$$

The next response y_{n+1} is observed at x_{n+1} and $\hat{\theta}_{n+1}$ is defined similarly. Since the $\{y_n\}$ are dependent, it is not obvious that standard results in linear model theory still hold. Three major issues to be studied are:

(A) Consistency of $\hat{\theta}_n$: Does $\hat{\theta}_n \to \theta$ with probability 1?

(B) Asymptotic distribution of $\hat{\theta}_n$: Does $(\hat{\theta}_n - \theta)^T M_n (\hat{\theta}_n - \theta) \to \sigma^2 \chi_p^2$ in distribution?

Consistency of $\hat{\sigma}^2$: $\hat{\sigma}^2 = \Sigma(y_i - x_i^T \hat{\theta}_n)^2/(n-p) \to \sigma^2$ with probability 1?

(C) Convergence of $n^{-1}M_n$ to an optimal design: Does $n^{-1}M_n \to D_\theta^*$, where $D_\theta^* = D(\eta^*)$ is an optimal design minimizing $\phi(g'(\theta)^T D(\eta) g'(\theta))$ over the normalized information matrix

$$D(\eta) = \int_X xx^T \eta(dx), \int_X \eta(dx) = 1?$$

Note that (A) implies the consistency of $\hat{\psi}_n$ to ψ and (B) implies the asymptotic validity of the standard confidence ellipsoid for θ, where F_α is the upper α point of the F distribution:

$$\{\theta : (\hat{\theta}_n - \theta)^T M_n (\hat{\theta}_n - \theta) \hat{\sigma}^{-2} < p^{-1}(n-p) F_\alpha(p, n-p)\} \quad .$$

The interpretation of (C) will be given for a special case. Take the optimality criterion to be the average of the asymptotic variances of the components of $\hat{\psi}_n$, that is, ϕ in (3.1) is the trace of a matrix. (C) says that the average variance of $\hat{\psi}_n$ for the design $\{x_1,\ldots,x_n\}$ is minimized as $n \to \infty$.

Questions (A) and (B) will be studied for more general sequential generation rules. Let x_{n+1} be an arbitrary measurable function of the past, $(x_1,y_1,\ldots,x_n,y_n)$. We assume that, for all i,

$$E(\varepsilon_i \mid \varepsilon_1,\ldots,\varepsilon_{i-1}) = 0,\ E(\varepsilon_i^2 \mid \varepsilon_1,\ldots,\varepsilon_{i-1}) = \sigma^2 < \infty \quad , \tag{3.3}$$

that is, ε_i is a martingale difference sequence with variance σ^2. We also assume that for some $\delta > 0$, with probability 1

$$\{\log \lambda_{max}(n)\}^{1+\delta}/\lambda_{min}(n) \to 0 \quad , \tag{3.4}$$

where $\lambda_{min}(n)$ and $\lambda_{max}(n)$ are the minimum and maximum eigenvalues of the random matrix M_n. Property (3.4) implies $\lambda_{min}(n) \to \infty$. Under (3.3) - (3.4), the strong consistency of $\hat{\theta}_n$ to θ follows from Corollary 3 of Lai and Wei (1982). This answers (A).

Before studying (B), we point out an underlying martingale structure that explains why standard asymptotic results for the fixed design problem hold for the sequential design problem under consideration. In $\hat{\theta}_n - \theta = M_n^{-1}(x_1\varepsilon_1 + \ldots + x_n\varepsilon_n)$, $\Sigma\, x_i\varepsilon_i$ is a martingale since x_i is a function of the past and ε_i is a martingale difference sequence. With the imposition of the growth rate condition (3.4) on x_i, the consistency of $\hat{\theta}_n$ follows from a martingale strong law of numbers. For the asymptotic normality of $\hat{\theta}_n$, the following stability condition on the random matrix M_n: there exists a non-random positive definite matrix B_n such that

$$B_n^{-1}(M_n)^{1/2} \to I_p \quad \text{and} \quad \max_i x_i^T B_n^{-2} x_i \to 0 \text{ in probability,} \tag{3.5}$$

ensures that $\hat{\theta}_n - \theta$ can be approximated by $B_n^{-2} \Sigma x_i \epsilon_i$, whose asymptotic normality follows from a martingale central limit theorem. Note that the stability condition (3.5) is considerably weaker than the objective in (C) that M_n/n converges to an optimal design matrix.

Under (3.3) and (3.5), the asymptotic normality of $\hat{\theta}_n$, i.e. $(\hat{\theta}_n-\theta)^T M_n(\hat{\theta}_n-\theta) \to \sigma^2\chi_p^2$, follows from Theorem 3 of Lai & Wei (1982). Under (3.3), the strong consistency of $\hat{\sigma}^2$ in (B) follows from Lemma 3 of Lai & Wei, whose only regularity condition, $n^{-1}\log \lambda_{max}(n) \to 0$ is satisfied since the design region is assumed bounded. Therefore, the standard confidence ellipsoid for θ is asymptotically valid under (3.5). The validity of the confidence region for ψ obtained from the confidence ellipsoid for θ by the g transformation needs the additional condition (3.4), which ensures the consistency of $\hat{\theta}_n$. A confidence ellipsoid for ψ can be constructed directly as

$$\{\psi : (\hat{\psi}_n-\psi)^T[g'(\hat{\theta}_n)^T M_n^{-1} g'(\hat{\theta}_n)]^{-1}(\hat{\psi}_n-\psi)\hat{\sigma}^{-2} < q^{-1}(n-p)F_\alpha(q,n-p)\} \quad .$$

Its asymptotic validity can be established from (A) and (B) under (3.3) - (3.5) as before.

We have answered questions (A) and (B) for very general rules that satisfy (3.3) - (3.5), which are, however, not easy to verify. For the simple example of §2, these conditions are either satisfied or not required. For general problems further discussions are given later in connection with (C).

We now consider question (C). If the normalized matrix $n^{-1}M_n$ converges to a nonsingular optimal design matrix D^*, it ensures that the conditions (3.4) - (3.5), required for (A) and (B), are satisfied.

The updating of M_n is governed by

$$(n+1)^{-1}M_{n+1} = \{1-(n+1)^{-1}\}n^{-1}M_n + (n+1)^{-1}x_{n+1}x_{n+1}^T \quad , \qquad (3.6)$$

where x_{n+1} is chosen according to (3.1) or (3.2), which depends on the current estimate $\hat{\theta}_n$. In the case where the criterion (3.1) or (3.2) is evaluated at the true parameter θ, the algorithm (3.6) has been studied extensively and the convergence of $n^{-1}M_n$ to D_θ^* was established for ϕ = determinant (Wynn, 1972; Pazman, 1974) and ϕ = trace (Wu & Wynn, 1978), assuming that D_θ^* is nonsingular. By a continuity argument, if $\hat{\theta}_n$ converges to θ with probability 1, then $n^{-1}M_n$ in (3.1) or (3.2) converges to D_θ^* with probability 1, for the same criterion. Since D_θ^* is assumed nonsingular, the above result does not cover the example in §2 with $|g(\theta)| = \frac{1}{2}$.

The strong consistency of $\hat{\theta}_n$, essential for the above argument, depends on the growth rate condition (3.4) on the random matrix M_n, which is not automatically satisfied by the selection rules (3.1) or (3.2). To ensure (3.4), the rules have to be modified so that the minimum eigenvalue of M_n grows to infinity at a rate no less than $(\log n)^{1+\delta}$ for some $\delta > 0$. That means, occasionally, we have to switch from (3.1) or (3.2) to a rule that maximizes the minimum eigenvalue of the augmented design matrix. The strong consistency of $\hat{\theta}_n$ is then guaranteed. It would be interesting to see if the convergence results for $n^{-1}M_n$ cited above still hold for the modified rules. It would then imply (3.5) and the asymptotic validity of the standard confidence regions. That is, (A), (B) and (C) would all be satisfied for such rules.

Appendix

Proof of (2.8)

From (2.2), the left-hand side of (2.8) can be approximated by

$$a_1(\eta_\theta^*(1))^{-1/2}(n-s_n)^{-1/2}\sum_{j=1}^{n-s_n}(y_{j+}-\phi_1) + a_2(\eta_\theta^*(-1))^{-1/2}s_n^{-1/2}\sum_{j=1}^{s_n}(y_{j-}-\phi_2) \quad . \tag{A.1}$$

Since, given s_n, $y_{j+} - \phi_1$ and $y_{j-} - \phi_2$ are independent and identically distributed, (A.1) with s_n replaced by $c_n = n\eta_\theta^*(-1)$ converges to the right-hand side of (2.8) from the central limit theorem. It remains to prove that in probability,

$$s_n^{-1/2}\sum_{j=1}^{s_n}(y_{j-}-\phi_2) - c_n^{-1/2}\sum_{j=1}^{c_n}(y_{j-}-\phi_2) \to 0 \quad , \tag{A.2}$$

and a similar expression for $n - s_n$. (A.2) follows easily from $s_n/c_n \to 1$ and the boundedness in probability of

$$s_n^{-1/2}\sum_{j=1}^{s_n}(y_{j-}-\phi_2) \quad \text{and} \quad (c_n-s_n)^{-1/2}\sum_{j=s_n+1}^{c_n}(y_{j-}-\phi_2) \quad .$$

REFERENCES

Box, G. E. P. and Hunter, W. G. (1965). Sequential design of experiments for nonlinear models. Proc. IBM Sc. Comp. Symp., 113-37, New York.

Cox, D. R. (1982). A remark on randomization in clinical trials. Utilitas Mathematica 21 A, 245-52.

Ford, I. and Silvey, S. D. (1980). A sequentially constructed design for estimating a nonlinear parametric function. Biometrika 67, 381-8.

Lai, T. L. and Wei, C. Z. (1982). Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. Ann. Statist. 10, 154-66.

Pazman, A. (1974). A convergence theorem in the theory of D-optimum experimental designs. Ann. Statist. 2, 216-8.

Siegmund, D. (1980). Sequential χ^2 and F tests and the related confidence intervals. Biometrika 67, 389-402.

Silvey, S. D. (1980). Optimal Design. London: Chapman and Hall.

Wu, C. F. and Wynn, H. P. (1978). The convergence of general step-length algorithms for regular optimum design criteria. Ann. Statist. 6, 1273-85.

Wynn, H. P. (1972). Results in the theory and construction of D-optimum experimental designs. J. R. Statist. Soc. B 34, 133-47.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 2853	2. GOVT ACCESSION NO. AD-A160 970	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ASYMPTOTIC INFERENCE FROM SEQUENTIAL DESIGN IN A NONLINEAR SITUATION		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) C. F. J. Wu		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53705		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE August 1985
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Martingale
Nonlinear design
Optimal design
Repeated-sampling principle
Sequential design

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

It is shown that confidence regions constructed by the repeated-sampling principle are asymptotically valid for sequential designs in general linear models and nonlinear parameters. The related questions of consistency of parameter estimators and convergence of sequential design to an optimal design are answered positively. An empirical finding of Ford and Silvey (1980) is given a theoretical justification.

END

FILMED

12-85

DTIC