

AN ASYMPTOTIC THEORY OF SEQUENTIAL DESIGNS BASED ON MAXIMUM LIKELIHOOD RECURSIONS

Zhiliang Ying and C. F. Jeff Wu

Rutgers University and University of Michigan

Abstract: Maximum likelihood recursions were proposed in Wu (1985, 1986) to obtain recursive procedures for nonlinear sequential design problems associated with many commonly used generalized linear models. It was argued empirically and heuristically there that these recursions should lead to asymptotically consistent and efficient designs. We prove that such recursions are consistent and asymptotically normal, at least for the location models including logistic, Poisson, gamma and inverse Gaussian. We show that a simple truncation leads to robust designs so that even if the models are incorrectly specified, the recursions still converge to the desired optimal design points. Asymptotic results concerning the sequential designs for the location-scale models are also obtained.

Key words and phrases: Asymptotic efficiency, asymptotic normality, generalized linear models, maximum likelihood estimator, Stochastic approximation, strong consistency.

1. Introduction

The stochastic approximation of Robbins and Monro (1951) is a sequential design for locating the zero of an unknown regression function. Let x denote the design point and y the corresponding response, whose mean M is a function of x . Robbins and Monro (1951) proposed to use x_n , which is generated from the following recursion

$$x_{n+1} = x_n - a_n y_n, \quad (1.1)$$

where a_n is a sequence of preassigned constants, to approximate the root M . They showed that, with a_n properly chosen, the sequentially determined x_n converges to the root of M . Numerous further refinements have been developed since the pioneering work of Robbins and Monro (cf. Blum (1954), Chung (1954), Sacks (1958), Robbins and Siegmund (1971) and Lai and Robbins (1979)). The stochastic approximation method has many important applications, including those in engineering: (Goodwin, Ramadge and Caines (1981), Kumar (1985)), biomedical science: (Finney (1978)) and educational testing: (Lord (1971)).

The Robbins-Monro recursion (1.1) can be interpreted as a maximum likelihood (ML) recursion when y and x follow the standard regression model $y =$

$\alpha + \beta x + \epsilon$, where ϵ is normally distributed with mean 0 and variance σ^2 . By solving the linear equation $\alpha + \beta x = 0$, the root M is $-\alpha/\beta$. By further assuming β is known, the ML estimate of M based on n observations $y_i, x_i, i = 1, \dots, n$ is

$$M_n = -\frac{\hat{\alpha}_n}{\beta} = \bar{x}_n - \frac{\bar{y}_n}{\beta}. \quad (1.2)$$

If the next observation y_{n+1} is taken at the current estimate M_n of M , i.e.,

$$x_{n+1} = M_n, \quad (1.3)$$

Lai and Robbins (1979) showed that (1.3) is equivalent to the Robbins-Monro recursion (1.1) with $a_n = (n\beta)^{-1}$. If β is unknown and estimated by the ML estimate $\hat{\beta}_n$ (which is also the least squares estimate), the ML recursion as defined in (1.3) with β replaced by $\hat{\beta}_n$ was interpreted by Wu (1986), eq. (11) as a special case of (1.1) with a_n taking a complicated form,

$$a_n = (n\hat{\beta}_n)^{-1} \left[1 + \frac{(n-1)^2 (\bar{y}_{n-1}/\hat{\beta}_{n-1})^2}{n \sum_{i=1}^n (x_i - \bar{x}_n)^2} \right]. \quad (1.4)$$

The observations made in these two papers are quite significant because they connect two seemingly distinct approaches to nonlinear sequential design. The approach that leads to (1.3) is *parametric* in that it is motivated by a linear function that links $E(y)$ and x and, to a lesser extent, by the normality of errors (which makes the least squares estimator fully efficient). On the other hand, the stochastic approximation (1.1) is *nonparametric* in that its asymptotic performance does not depend on the knowledge of $E(y)$ as a function of x . The assumption $y = \alpha + \beta x + \epsilon$ is useful for motivating and generating design procedures. The validity and performance of the resulting design are nonetheless independent of the assumption.

Once this connection is recognized, we can greatly expand both approaches to cover more general variations. From the likelihood point of view, the Robbins-Monro recursion has better efficiency when the errors are normal or nearly normal, even though it is consistent and asymptotically normal for very general error distributions. For distributions that are distinctly different from normal, e.g., binomial, Poisson, gamma and inverse Gaussian, it is more natural to extend (1.1) by adopting the ML recursion approach with the likelihood or quasi-likelihood function capturing the nature of the variation. This is essentially the viewpoint taken in Wu (1986). To illustrate this approach, consider a binary experiment in which the outcome y is denoted by 1 (response) or 0 (nonresponse). The probability of response is related to a stress level x (at which the experimentation is performed) by

$$G(x) = \text{Prob}\{y = 1|x\} = E(y|x).$$

Suppose the interest focuses on estimating the 100p-th percentile α_p of $G(x)$, i.e., $G(\alpha_p) = p$. Since G is unknown, we assume (correctly or not) that G follows a one-parameter model $H(x - \alpha)$ with H known, $H(0) = p$ and α to be estimated. Under this assumption we can estimate the only parameter α in the model by using the ML method and denote it by $\hat{\alpha}_n$. Then we can choose the next design point x_{n+1} to be the current best estimate of α_p , which is $\hat{\alpha}_n$ under the given model assumption. So, in the ML recursion design

$$x_{n+1} = \hat{\alpha}_n.$$

Further discussions on this recursion scheme, its relation to the Robbins-Monro recursion (1.1), and extensions to cover the location-scale model $H(\beta(x - \alpha))$ can be found in Wu (1985, 1986) and the later sections of the present paper. Extensions of this recursive scheme to generalized linear models were given in Wu (1986). Wu's description of a sequential design of Poisson experiments is of particular interest.

In general, when the underlying probability law of y given x is specified up to a finite number of parameters, Wu (1986) proposed using, at each stage, the updated maximum likelihood estimate to set the next design point. To be specific, suppose the density of the response y given x is $f(\cdot|x, \theta)$, where θ is an unknown parameter vector, and the objective of the design is to select x so that $E(y|x, \theta)$ stays as closely as possible to some preset target value p . Let $g(\theta)$ denote the unique value determined from

$$E(y|x = g(\theta), \theta) = p.$$

Wu's proposal is to compute the ML estimator $\hat{\theta}_n$ based on the first n observations and set the next design point $x_{n+1} = g(\hat{\theta}_n)$, which is the best current "guess" of the target value $g(\theta)$. Because of its full and efficient use of the data, this procedure is likely to bring x_n close to $g(\theta)$ in a relatively fewer number of steps. Indeed, for the binary response data, Wu (1985) has demonstrated, both empirically and heuristically, the advantages of using the ML recursion for the sequential design problem.

In this paper we are concerned with the convergence of the sequential designs derived from the maximum likelihood and its related recursions. In the next section, we show that the ML recursions for one-parameter location models lead to consistent and asymptotically normal design sequences under certain regularity conditions, which are verifiable for the commonly used generalized linear models. A simple truncation is introduced in Section 3, where it is shown that with such modification, consistency and asymptotic normality still hold even if the model is incorrectly specified. Section 4 presents similar asymptotic results for the sequential designs in the two-parameter location-scale models.

2. Sequential Designs Based on Location Models

Here we consider a generalized linear model: if the design level is set at x , the mean response of y is given by

$$E(y|x) = H(x - \alpha), \quad (2.1)$$

where H is a known function and α an unknown parameter. Without loss of generality, let $p = H(0)$ be the value of the desired mean response for y . It is known that in many situations the optimal design for minimizing the asymptotic variance of the ML estimator of α is to put all the design points at $x = \alpha$. Wu (1988) proved this for the binomial variation and Ford, Torsney and Wu (1992) and Sitter and Torsney (1995) extended it to generalized linear models. Assume the variance of y at x is of the form $V(x - \alpha)$. The setup includes all location models for which the effect of the design is a shift of location on the distribution of y , or any parametric models that can be transformed into the location models.

We shall study the asymptotic behavior of $\{x_n\}$, a random sequence with x_1 the initial value and the subsequent x 's defined recursively by

$$\sum_{i=1}^n \psi(x_i) [y_i - H(x_i - x_{n+1})] = 0, \quad n = 1, 2, \dots, \quad (2.2)$$

where $\psi \geq 0$ is a prespecified weight function. The recursion actually combines two steps involving estimation of α and design of x . Letting $\hat{\alpha}_n$ denote x_{n+1} , (2.2) can alternately be expressed as

$$\sum_{i=1}^n \psi(x_i) [y_i - H(x_i - \hat{\alpha}_n)] = 0, \quad (2.3)$$

$$x_{n+1} = \hat{\alpha}_n. \quad (2.4)$$

The preceding recursion was proposed by Wu (1986) and is essentially a myopic strategy for approximating the optimal design $x = \alpha$ in the ideal situation of known location parameter α . The first step (2.3) mimics the maximum likelihood estimating equation for α . In fact, with ψ properly chosen, it becomes the maximum likelihood estimating equation for the five most commonly used generalized linear models as given in Table 2.1 of McCullagh and Nelder (1989). Equation (2.4) is just the most obvious myopic way of selecting the best current design level.

It is not clear how to prove the convergence of the design sequence $\{x_n\}$ if it is at all true. The difficulty lies in the complicated dependency among the x_n and y_n . Furthermore, because (2.2) involves the whole history of x_i , $i = 1, \dots, n$, the techniques developed for proving convergence of the classical

stochastic approximation algorithms, which connect x_{n+1} to y_n and x_n only, are not applicable. We adopt a different approach here. For technical reasons, we need to introduce the following conditions:

- (C1) H is continuous and strictly increasing, and with probability 1; (2.2) is well defined for all large n .
- (C2) for every $K > 0$, $\infty > \sup_{|t| \leq K} \{\psi(t+\alpha)V(t)\} \geq \inf_{|t| \leq K} \{\psi(t+\alpha)V(t)\} > 0$.
- (C3) $\liminf_{|t| \rightarrow \infty} \frac{|H(t)-p|}{V(t)\psi(t+\alpha)} > 0$ and $\liminf_{|t| \rightarrow \infty} \frac{|H(2t)-H(t)|}{V(t)\psi(t+\alpha)} > 0$.
- (C4) V is continuous at 0, H is continuously differentiable in a neighborhood of 0 and $H'(0) > 0$.

Conditions (C1), (C2) and (C4) are satisfied by almost all sensible models. Condition (C3) is more restrictive, but is verifiable in the subsequent examples, which motivate this investigation. It is about the tail growth rates of the mean, variance and weight functions. We now use the five generalized linear models listed in Table 2.1 of McCullagh and Nelder (1989) to illustrate the sequential design given by (2.2) and conditions (C1)–(C4).

Example 1. (Normal model) Suppose that the distribution of y given x is $N(x - \alpha, \sigma^2)$, where σ^2 may or may not be known. Given observations $x_1, y_1, \dots, x_n, y_n$ (at stage n), the ML estimating equation for α is

$$\sum_{i=1}^n [y_i - (x_i - \hat{\alpha}_n)] = 0. \tag{2.5}$$

Thus $\phi \equiv 1$, $H(t) = t$ and $V(t) = \sigma^2$. Conditions (C1)–(C4) are clearly satisfied. From (2.5) we get

$$x_{n+1} = \hat{\alpha}_n = -\frac{1}{n} \sum_{i=1}^n (y_i - x_i) = \alpha - \frac{1}{n} \sum_{i=1}^n \epsilon_i,$$

where $\epsilon_i = y_i - (x_i - \alpha)$ are i.i.d. $N(0, \sigma^2)$.

Example 2. (Logit model for binary response data) As before, let p denote the target mean response, which must be strictly between 0 and 1. Then the mean, variance and weight functions are respectively

$$H(t) = \frac{pe^t}{1-p+pe^t}, \quad V(t) = \frac{p(1-p)e^t}{(1-p+pe^t)^2} \quad \text{and} \quad \psi(t) = 1.$$

So the “success” probability of y given x is $H(x-\alpha)$. The ML estimating equation of α at stage n is

$$\sum_{i=1}^n \left(y_i - \frac{pe^{x_i - \hat{\alpha}_n}}{1-p+pe^{x_i - \hat{\alpha}_n}} \right) = 0. \tag{2.6}$$

It is easy to see that if the y_i in (2.6) take both 0 and 1 values, then $\hat{\alpha}_n$ is uniquely defined. Furthermore, (C1), (C2) and (C4) are clearly satisfied, and (C3) can be verified directly.

Example 3. (Poisson model) In this case, $P(y=k|x) = (pe^{x-\alpha})^k \exp\{-pe^{x-\alpha}\}/k!$ and the ML estimating equation of α at stage n is

$$\sum_{i=1}^n (y_i - pe^{x_i - \hat{\alpha}_n}) = 0. \quad (2.7)$$

So $V(t) = H(t) = pe^t$ and $\psi(t) \equiv 1$. Also, as long as at least one of y_i , $i \leq n$ is not 0, (2.7) has a unique solution $x_{n+1} = \hat{\alpha}_n$. Conditions (C1)–(C4) can be easily verified.

Example 4. (Gamma model) Following McCullagh and Nelder (1989), we assume, for the gamma model, that the density function y given x is

$$f(u|x, \alpha, \nu) = \exp\{-\nu[p^{-1}e^{-(x-\alpha)}u + x - \alpha] - \nu \log p + (\nu-1) \log u - \log \Gamma(\nu) + \nu \log \nu\},$$

where ν is the shape parameter that is not required to be known. It follows that $H(t) = pe^t$ and $V(t) = \nu^{-1}p^2e^{2t}$. Furthermore, the ML estimating equation becomes

$$\sum_{i=1}^n (y_i - pe^{x_i - \hat{\alpha}_n})e^{-x_i} = 0, \quad (2.8)$$

so $\psi(t) = e^{-t}$. Equation (2.8) always has a unique solution except when all y_i 's are 0, which has probability 0. Again (C1), (C2) and (C4) are clearly satisfied. Note that $|H(t) - p|/[V(t)\psi(t)] = \nu p^{-1}|e^t - 1|e^{-t} = \nu p^{-1}|1 - e^{-t}|$ and $|H(2t) - H(t)|/[V(t)\psi(t)] = \nu p^{-1}|e^{2t} - e^t|e^{-t} = \nu p^{-1}|e^t - 1|$. So (C3) is also satisfied.

Example 5. (Inverse Gaussian) Here we assume the density function of y given x to be

$$f(u|x, \alpha, \sigma^2) = \exp\{[-2^{-1}p^{-2}e^{-2(x-\alpha)}u + p^{-1}e^{-(x-\alpha)}]/\sigma^2 + g(\sigma^2, u)\},$$

where σ^2 may or may not be known, $g(\sigma^2, u) = -2^{-1}[\log(2\pi\sigma^2u^3) + (\sigma^2u)^{-1}]$, a function not involving α . Thus we have $H(t) = pe^t$ and $V(t) = \sigma^2p^3e^{3t}$. The ML estimating equation for α is

$$\sum_{i=1}^n e^{-2x_i}(y_i - pe^{x_i - \hat{\alpha}_n}) = 0. \quad (2.9)$$

The weight function is $\psi(t) = e^{-2t}$. As before, (C1), (C2) and (C4) hold trivially and (C3) can be easily verified.

There are other situations in which the ML estimating equations have to be modified in order to make use of (2.2). For example, in probit analysis, the response y given x is assumed to be Bernoulli with success probability $\Phi(x - \alpha + c_p)$, where Φ is the normal curve and c_p its p th quantile. Letting ϕ be the normal density function, the ML estimator $\hat{\alpha}_n^*$ solves

$$\sum_{i=1}^n \frac{\phi(x_i - \hat{\alpha}_n^* + c_p)}{\Phi(x_i - \hat{\alpha}_n^* + c_p)(1 - \Phi(x_i - \hat{\alpha}_n^* + c_p))} [y_i - \Phi(x_i - \hat{\alpha}_n^* + c_p)] = 0. \quad (2.10)$$

This equation is not a special case of (2.3) because the weight function depends intrinsically on α . However, since when x_n is close to α , the corresponding weight becomes close to some constant, we can modify (2.10) to

$$\sum_{i=1}^n [y_i - \Phi(x_i - \hat{\alpha}_n + c_p)] = 0, \quad (2.11)$$

which is clearly a special case of (2.2). Moreover, conditions (C1), (C2) and (C4) are trivially satisfied; to check (C3), we note that $\lim_{|t| \rightarrow \infty} (|\Phi(t) - p|) / (\Phi(t)(1 - \Phi(t))) = \infty$ and

$$\lim_{t \rightarrow \infty} \frac{\Phi(2t - c_p) - \Phi(t)}{\Phi(t)(1 - \Phi(t))} = - \lim_{t \rightarrow -\infty} \frac{\Phi(2t - c_p) - \Phi(t)}{\Phi(t)(1 - \Phi(t))} = \lim_{t \rightarrow \infty} \left(1 - \frac{1 - \Phi(2t - c_p)}{1 - \Phi(t)} \right) = 1.$$

The next theorem contains the main results of the section and is applicable to all the preceding examples. It shows that under suitable conditions, the sequential designs of the form (2.2) are strongly consistent and asymptotically normal.

Theorem 1. *Let y_i , x_i be defined as in the beginning of the section. Suppose the corresponding mean, variance and weight functions H , V and ψ satisfy conditions (C1)-(C3). Define the disjoint events $A_\alpha = \{x_n \rightarrow \alpha\}$, $A_\infty = \{x_n \rightarrow \infty\}$ and $A_{-\infty} = \{x_n \rightarrow -\infty\}$.*

- (i) $P\{A_\alpha \cup A_\infty \cup A_{-\infty}\} = 1$; in fact $P\{A_\alpha | \sum_{n=1}^{\infty} \psi^2(x_n)V(x_n - \alpha) = \infty\} = 1$ and $P\{A_\infty \cup A_{-\infty} | \sum_{n=1}^{\infty} \psi^2(x_n)V(x_n - \alpha) < \infty\} = 1$.
- (ii) If $\liminf_{t \rightarrow \infty} \psi^2(t)[1 + V(t - \alpha)] > 0$, then $P(A_\infty) = 0$; and if $\liminf_{t \rightarrow -\infty} \psi^2(t)[1 + V(t - \alpha)] > 0$, then $P(A_{-\infty}) = 0$.
- (iii) Suppose V and ψ satisfy the tail growth condition

$$\liminf_{|t| \rightarrow \infty} \psi^2(t)[1 + V(t - \alpha)] > 0. \quad (2.12)$$

Then $x_n \rightarrow \alpha$ a.s.

- (iv) If (C4) is also satisfied and $x_n \rightarrow \alpha$ a.s., then $\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N(0, (H'(0))^{-2}V(0))$. In particular, if (C4) and (2.12) hold, then $\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N(0, (H'(0))^{-2}V(0))$.

Remark 1. Part (i) implies that $x_n \rightarrow \alpha$ if and only if $\sum_{n=1}^{\infty} \psi^2(x_n)V(x_n - \alpha) = \infty$.

Remark 2. We have shown that the sequential designs derived from the ML recursions in Examples 1–5 satisfy conditions (C1)–(C4). Thus, in view of Theorem 1(iv), in order to show any of these designs to be strongly consistent, it suffices to verify that the corresponding ψ and V satisfy (2.12). For $\psi \equiv 1$, (2.12) holds trivially. So, for Example 1, the normal model,

$$\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N(0, \sigma^2);$$

for Example 2, the logit model,

$$\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N\left(0, [p(1-p)]^{-1}\right);$$

for Example 3, the Poisson model,

$$\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N(0, p^{-1});$$

for the probit model with modified maximum likelihood recursion (2.11),

$$\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N(0, 2\pi p(1-p)e^{\frac{c^2}{p}}).$$

Now for Example 4, the gamma model, $\psi^2(t)(1+V(t)) \geq \nu^{-1}p^2 > 0$. So (2.12) is satisfied and the sequential design x_n satisfies

$$\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N(0, \nu^{-1}).$$

Remark 3. Theorem 1(i) and (ii) are useful for situations in which (2.12) either does not hold or is difficult to verify. We use Example 5 to illustrate this. In this case, we have $\liminf_{t \rightarrow \infty} \psi^2(t)[1+V(t-\alpha)] = 0$, but $\liminf_{t \rightarrow -\infty} \psi^2(t)[1+V(t-\alpha)] = \infty$. Therefore, (iii) is not applicable but (ii) implies that $P(x_n \rightarrow -\infty) = 0$. In view of (i), in order to show $x_n \rightarrow \alpha$ a.s., it suffices to rule out $x_n \rightarrow \infty$. Again by (i), $x_n \rightarrow \infty$ can only happen on $\{\sum_{n=1}^{\infty} \psi^2(x_n)V(x_n - \alpha) < \infty\}$, which implies the convergence of $\sum_{i=1}^n (y_i - pe^{x_i - \alpha})\psi_i$. But by definition $\sum_{n=1}^n y_i\psi_i = \sum_{n=1}^n pe^{x_i - x_{n+1}}$. So $\sum_{n=1}^n [pe^{x_i - x_{n+1}} - pe^{x_i - \alpha}]\psi_i = p(e^{-x_{n+1}} - e^{-\alpha})\sum_{n=1}^n e^{x_i}$ converges to a finite limit. But this obviously rules out $x_n \rightarrow \infty$. So we conclude $x_n \rightarrow \alpha$ a.s. and

$$\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N(0, \sigma^2 p).$$

Proof of Theorem 1. We first note that (iii) is a direct consequence of (i) and (ii). Furthermore, (iv) is also easy to get via the following argument. From (2.2),

$$\sum_{i=1}^n \psi(x_i) [y_i - H(x_i - \alpha)] = \sum_{i=1}^n \psi(x_i) [H(x_i - x_{n+1}) - H(x_i - \alpha)]. \quad (2.13)$$

Since $x_n \rightarrow \alpha$, we can apply Taylor's expansion to (2.13) to get

$$\sqrt{n}(x_{n+1}-\alpha) = - [\psi(\alpha)H'(0) + o_p(1)]^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n [y_i - H(x_i - \alpha)] \psi(x_i) + o_p(1),$$

which converges to $N(0, [H'(0)]^{-2}V(0))$ by the martingale central limit theorem (Pollard (1984), p. 171).

It remains to prove (i) and (ii). A key idea in the subsequent proof is to use the martingale local convergence theorem of Chow (1965), Corollary 5, from which it follows that

$$\sum_{n=1}^{\infty} \frac{\psi(x_n) [y_n - H(x_n - \alpha)]}{\sum_{i=1}^n \psi^2(x_i)V(x_i - \alpha)} \text{ converges a.s.} \quad (2.14)$$

On the set that the denominator in (2.14) goes to infinity, it will be argued via Kronecker's lemma that the design sequence converges to α , whereas on its complement, it goes to either ∞ or $-\infty$.

We first prove that on the event $\{\sum_{n=1}^{\infty} \psi^2(x_n)V(x_n - \alpha) < \infty\}$, $x_n \rightarrow \infty$ or $x_n \rightarrow -\infty$ a.s. On this event we have $\psi^2(x_n)V(x_n - \alpha) \rightarrow 0$, which implies $|x_n| \rightarrow \infty$ in view of (C2). Thus all we need to do is to rule out the possibility that $\limsup x_n = \infty$ and $\liminf x_n = -\infty$ occur simultaneously. If this were true, then, since $\lim |x_n| = \infty$, we could find a subsequence n_k such that $x_{n_k} \rightarrow -\infty$ and $x_{n_k+1} \rightarrow \infty$ as $k \rightarrow \infty$. From (2.2), we would have

$$\begin{aligned} \psi(x_{n_k}) [y_{n_k} - H(x_{n_k} - x_{n_k+1})] &= \sum_{i=1}^{n_k-1} [H(x_i - x_{n_k+1}) - y_i] \psi(x_i) \\ &= \sum_{i=1}^{n_k-1} [H(x_i - x_{n_k+1}) - H(x_i - x_{n_k})] \psi(x_i) \\ &\leq \sum_{i=1}^m [H(x_i - x_{n_k+1}) - H(x_i - x_{n_k})] \psi(x_i) \end{aligned}$$

for every fixed $m < n_k$ since $H(x_i - x_{n_k+1}) - H(x_i - x_{n_k}) \leq 0$. Equivalently,

$$\begin{aligned} \psi(x_{n_k}) [y_{n_k} - H(x_{n_k} - \alpha)] &\leq \sum_{i=1}^m [H(x_i - x_{n_k+1}) - H(x_i - x_{n_k})] \psi(x_i) \\ &\quad + \psi(x_{n_k}) [H(x_{n_k} - x_{n_k+1}) - H(x_{n_k} - \alpha)]. \end{aligned} \quad (2.15)$$

But the left-hand side of (2.15) $\rightarrow 0$ in view of (2.14), while the right-hand side of (2.15) can be made smaller than a fixed negative number. Thus the desired contradiction occurs.

We now prove $x_n \rightarrow \alpha$ on the event $\{\sum_{n=1}^{\infty} \psi^2(x_n)V(x_n - \alpha) = \infty\}$. By (2.14) and Kronecker's Lemma,

$$\frac{\sum_{i=1}^n \psi(x_i)[y_i - H(x_i - \alpha)]}{\sum_{i=1}^n \psi^2(x_i)V(x_i - \alpha)} \longrightarrow 0 \quad \text{a.s.} \quad (2.16)$$

Combining (2.16) with (2.2), we get

$$\frac{\sum_{i=1}^n \psi(x_i)[H(x_i - x_{n+1}) - H(x_i - \alpha)]}{\sum_{i=1}^n \psi^2(x_i)V(x_i - \alpha)} \longrightarrow 0 \quad \text{a.s.} \quad (2.17)$$

If $\sup_n |x_n| < \infty$, then because of the strict monotonicity of H , (2.17) implies $x_n \rightarrow \alpha$. So we only need to prove that $\sup_n |x_n| < \infty$ indeed holds. Suppose it did not. Then we could find a subsequence x_{n_k} such that $|x_{n_k}| \rightarrow \infty$ and $|x_i| \leq |x_{n_k}|$, $i \leq n_k$. By choosing a further subsequence, we may assume, without loss of generality, that $x_{n_k} \rightarrow \infty$.

From (C2) and (C3), there exist constants $\varepsilon_0 > 0$, $C_0 > 0$ such that

$$\frac{p - H(x_i - \alpha)}{V(x_i - \alpha)\psi(x_i)} \leq -\varepsilon_0 \quad \text{for all } x_i \geq C_0.$$

Since $x_i \leq x_{n_k}$, this shows

$$\frac{H(x_i - x_{n_k}) - H(x_i - \alpha)}{V(x_i - \alpha)\psi(x_i)} \leq \frac{p - H(x_i - \alpha)}{V(x_i - \alpha)\psi(x_i)} \leq -\varepsilon_0 \quad (2.18)$$

for all $x_i \geq C_0$. For $|x_i| \leq C_0$, it is easy to see

$$\frac{H(x_i - x_{n_k}) - H(x_i - \alpha)}{V(x_i - \alpha)\psi(x_i)} \leq -\varepsilon_1 \quad (2.19)$$

for some $\varepsilon_1 > 0$ and large k because of the strict monotonicity of H . For $x_i \leq -C_0$, since $H(x_i - x_{n_k}) = H(x_i - \alpha - (x_{n_k} - \alpha)) \leq H(2(x_i - \alpha))$,

$$\frac{H(x_i - x_{n_k}) - H(x_i - \alpha)}{V(x_i - \alpha)\psi(x_i)} \leq \frac{H(2(x_i - \alpha)) - H(x_i - \alpha)}{V(x_i - \alpha)\psi(x_i)} \leq -\varepsilon_2 \quad (2.20)$$

for some positive ε_2 , where the last inequality follows from (C3). In view of (2.18)–(2.20),

$$\limsup_{k \rightarrow \infty} \frac{\sum_{i=1}^{n_k} [H(x_i - x_{n_k}) - H(x_i - \alpha)]}{\sum_{i=1}^{n_k} V(x_i - \alpha)} \leq -\min(\varepsilon_0, \varepsilon_1, \varepsilon_2) < 0,$$

which clearly contradicts (2.17). Thus (i) holds.

For (ii), we only prove the first part, i.e., $\liminf_{t \rightarrow \infty} \psi^2(t)[1 + V(t - \alpha)] > 0$ implies $P(x_n \rightarrow \infty) = 0$. The second part can be proven in exactly the same

way. From the stated assumption, we have either $\liminf_{t \rightarrow \infty} \psi^2(t)V(t - \alpha) > 0$ or $\liminf_{t \rightarrow \infty} \psi^2(t) > 0$. If $\liminf_{t \rightarrow \infty} \psi^2(t)V(t - \alpha) > 0$, then on $\{x_n \rightarrow \infty\}$ we necessarily have $\sum_n \psi^2(x_n)V(x_n - \alpha) = \infty$. But by (i), $x_n \rightarrow \alpha$ a.s., which is impossible. Suppose now that $\liminf_{t \rightarrow \infty} \psi^2(t) > 0$. We prove that $x_n \rightarrow \infty$ is impossible by contradiction. On $\{x_n \rightarrow \infty\}$, $\sum_n \psi^2(x_n)V(x_n - \alpha) < \infty$ by (i). This and (2.14) imply that

$$\psi(x_n) [y_n - H(x_n - \alpha)] \longrightarrow 0. \quad (2.21)$$

On the other hand, from (2.2), with n_k chosen so that $x_{n_k} \leq x_{n_k+1}$,

$$\begin{aligned} & \psi(x_{n_k}) [y_{n_k} - H(x_{n_k} - \alpha)] \\ &= \sum_{i=1}^{n_k} \psi(x_i) [H(x_i - x_{n_k+1}) - H(x_i - x_{n_k})] + \psi(x_{n_k}) [p - H(x_{n_k} - \alpha)] \\ &\leq \psi(x_{n_k}) [p - H(x_{n_k} - \alpha)] \end{aligned} \quad (2.22)$$

Since (C2) entails $\psi > 0$, we have

$$y_{n_k} - H(x_{n_k} - \alpha) \leq p - H(x_{n_k} - \alpha) \quad (2.23)$$

from (2.22). The left-hand side of (2.23) converges to zero while its right-hand side converges to $p - H(\infty) < 0$, which is a contradiction.

3. Extension to Models with Misspecified Link(mean) Functions

The Robbins-Monro stochastic approximation is nonparametric in the sense that, except for a local monotonicity assumption, no prior knowledge of the mean response function is assumed. In contrast, the analysis of the ML-based sequential design in the preceding section hinges on the correct specification of the link function H . In this section, we deal with the situation in which the link function is misspecified. It will be shown that if the optimal level α is known to lie in a bounded interval, then a simple modification of (2.2) or, more precisely, (2.3)–(2.4) by truncation leads to a strongly consistent design sequence. Asymptotic normality for the design sequence is also obtained under certain regularity conditions.

As in Section 2, the response y at x has mean $H(x - \alpha)$ and variance $V(x - \alpha)$, where, throughout this section, H is assumed to be continuous, strictly increasing and $H(0) = p$, and V is assumed to be positive and continuous. However, \tilde{H} , another mean function, is used to generate the design sequence:

$$\sum_{i=1}^n \psi(x_i) (y_i - \tilde{H}(x_i - \hat{\alpha}_n)) = 0,$$

$$x_{n+1} = \max \{ \min(\hat{\alpha}_n, b), a \}, \quad (3.1)$$

where a and b are predetermined constants and ψ is bounded away from 0 and infinity in $[a, b]$. We need the following conditions for the analysis of (3.1):

- (C5) \tilde{H} is twice continuously differentiable, $\tilde{H}'(t) > 0$ and $\tilde{H}(0) = p$;
(C6) There exists $C > 0$ such that $\tilde{H}(-C) < H(x - \alpha) < \tilde{H}(C)$ for all $x \in [a, b]$;
(C7) $\sup_{t \in [a, b]} E[|y|^4 | x = t] < \infty$.

Theorem 2. *Under conditions (C5)–(C7), $x_n \rightarrow \alpha$ a.s.*

Proof. By the martingale local convergence theorem of Chow (1965) we know that (2.14) still holds. Since $a \leq x_n \leq b$, we have $\sum_{n=1}^{\infty} V(x_n - \alpha)\psi(x_n) = \infty$. Thus Kronecker's lemma implies

$$\frac{\sum_{i=1}^n [y_i - H(x_i - \alpha)]\psi(x_i)}{\sum_{i=1}^n V(x_i - \alpha)\psi(x_i)} \rightarrow 0 \quad \text{a.s.} \quad (3.2)$$

From (C6), there exist $\Delta > 0$ and $\delta > 0$ such that $\tilde{H}(x - \Delta) - H(x - \alpha) \leq -\delta$ and $\tilde{H}(x + \Delta) - H(x - \alpha) \geq \delta$ for all $x \in [a, b]$. So for any $t \geq \Delta$ (or $t \leq -\Delta$),

$$\sum_{i=1}^n [\tilde{H}(x_i - t) - H(x_i - \alpha)]\psi(x_i) \leq -\delta \sum_{i=1}^n \psi(x_i) \quad \left(\text{or } \geq \delta \sum_{i=1}^n \psi(x_i) \right).$$

But in view of the first equation in (3.1) and (3.2), we have

$$\frac{1}{n} \sum_{i=1}^n [\tilde{H}(x_i - \hat{\alpha}_n) - H(x_i - \alpha)]\psi(x_i) \rightarrow 0 \quad \text{a.s.},$$

implying that $|\hat{\alpha}_n| \leq \Delta$ for all large n . Thus, the truncation step in (3.1) is in effect only for a finite number of steps and eventually $x_{n+1} = \hat{\alpha}_n$. Furthermore, modifying the tails of \tilde{H} if necessary, we can assume that $\tilde{H}'(x)$ is bounded away from 0 and $|\tilde{H}''(x)|$ is bounded away from ∞ .

The main idea in the rest of our proof for the strong consistency is to first transform the sequential design (3.1) into a Robbins-Monro-type recursion and then apply a convergence result established for the latter procedure. From the definition of α ,

$$\psi(x_n)(y_n - p) = \sum_{i=1}^n \psi(x_i) [\tilde{H}(x_i - \hat{\alpha}_n) - \tilde{H}(x_i - \hat{\alpha}_{n-1})] = J_n(\hat{\alpha}_{n-1} - \hat{\alpha}_n), \quad (3.3)$$

where $J_n(u) = \sum_{i=1}^n \psi(x_i) [\tilde{H}(x_i - \hat{\alpha}_{n-1} + u) - \tilde{H}(x_i - \hat{\alpha}_{n-1})]$. Obviously J_n is strictly increasing, twice continuously differentiable and $J_n(0) = 0$. Inverting J_n in (3.3) we get

$$\hat{\alpha}_n = \hat{\alpha}_{n-1} - J_n^{-1}(\psi(x_n)(y_n - p)). \quad (3.4)$$

Letting $Q_n(v) = J_n^{-1}(v)$, we have

$$Q_n'(v) = \frac{1}{J_n'(Q_n(v))}, \quad Q_n''(v) = \frac{-J_n''(Q_n(v))}{J_n'^3(Q_n(v))}.$$

By taking a Taylor's expansion of Q_n at 0, we obtain

$$\hat{\alpha}_n = \hat{\alpha}_{n-1} - \frac{\psi(x_n)}{J_n'(0)}(y_n - p) + \xi_n, \quad (3.5)$$

where $\xi_n = J_n''(\theta_n^*)\psi^2(x_n)(y_n - p)^2/[2J_n'^3(\theta_n^*)]$ for some θ_n^* between 0 and $\hat{\alpha}_{n-1} - \hat{\alpha}_n$. Let \mathcal{F}_{n-1} be the σ -field generated by $y_i, x_i, i \leq n-1$. Since $|J_n''(\theta_n^*)/J_n'^3(\theta_n^*)| \leq \delta_0 n^{-2}$ for some $\delta_0 > 0$, we have $E(|\xi_n| + \xi_n^2 | \mathcal{F}_{n-1}) = O(n^{-2})$ a.s. in view of (C7). From this and (3.5) it follows that

$$E \left\{ (\hat{\alpha}_n - \alpha)^2 | \mathcal{F}_{n-1} \right\} \leq (\hat{\alpha}_{n-1} - \alpha)^2 - \frac{\psi(x_n)}{J_n'(0)} (\hat{\alpha}_{n-1} - \alpha) (H(x_n - \alpha) - p) + \eta_n \quad (3.6)$$

with $\eta_n \geq 0$ and $\sum_{n=1}^{\infty} \eta_n < \infty$. Therefore, Theorem 1 of Robbins and Siegmund (1971) can be applied to (3.6) to obtain that $\lim_{n \rightarrow \infty} (\hat{\alpha}_n - \alpha)^2$ exists and

$$\sum_{n=1}^{\infty} \frac{\psi(x_n)}{J_n'(0)} (\hat{\alpha}_{n-1} - \alpha) (H(x_n - \alpha) - p) < \infty.$$

Since H is strictly increasing at 0 and $x_n = \hat{\alpha}_{n-1}$ for all large n , we must have $\lim_{n \rightarrow \infty} \hat{\alpha}_n = \alpha$ a.s. or equivalently $\lim_{n \rightarrow \infty} x_n = \alpha$ a.s.

Theorem 3. *Suppose that the assumptions in Theorem 2 are satisfied, ψ is continuous at α and $\tilde{H}'(0) < 2H'(0)$. Then*

$$\sqrt{n} (x_n - \alpha) \xrightarrow{\mathcal{L}} N \left(0, \frac{V(0)}{\tilde{H}'(0)[2H'(0) - \tilde{H}'(0)]} \right).$$

Proof. The idea here is to show that the procedure is asymptotically equivalent to the Robbins-Monro recursion so that the asymptotic normality as given in Lai and Robbins (1979) for the latter can be applied to the former. From Theorem 2, $x_{n+1} = \hat{\alpha}_n$ for all large n . Thus we can deduce from (3.5) that

$$\begin{aligned} x_{n+1} &= x_n - \left[\frac{\psi(x_n)}{J_n'(0)} - \frac{J_n''(\theta_n^*)}{2J_n'^3(\theta_n^*)} \psi^2(x_n)(y_n - p) \right] (y_n - p) \\ &= x_n - \frac{1}{n\beta_n^*} (y_n - p), \text{ say,} \end{aligned}$$

for some θ_n^* between 0 and $x_n - x_{n+1}$. Clearly

$$\beta_n^* - \frac{J_n'(0)}{n\psi(x_n)} = O \left(\frac{y_n - p}{n} \right) = o_p(n^{-\frac{3}{5}}),$$

and $J'_n(0)/[n\psi(x_n)] \in \mathcal{F}_{n-1}$ and converges to $\tilde{H}'(0)$ a.s.. Therefore, Theorem 4 (iv) of Lai and Robbins (1979) can be used to obtain the desired convergence for $\sqrt{n}(x_n - \alpha)$. Note that the i.i.d. assumption on $y_i - H(x_i - \alpha)$ in Lai and Robbins can be relaxed to cover the current setup.

4. Sequential Designs in Location-Scale Models with Estimated Scale Parameter

When the link function H in (2.1) is specified only up to a scale parameter, we are faced with a more general model:

$$E(y|x) = H(\beta(x - \alpha)), \quad (4.1)$$

where both α and $\beta > 0$ are unknown. Wu (1985) gave an extensive analysis in the case of a logistic link function for which $E(y|x) = e^{\beta(x-\alpha)}/[1 + e^{\beta(x-\alpha)}]$. Lai and Robbins (1979) in their construction of optimal adaptive stochastic approximation algorithms obtained strong consistency and asymptotic normality for the design sequence with the scale parameter estimated sequentially. Following Lai and Robbins, we first assume that a strongly consistent sequence of estimators $\hat{\beta}_n$ is available. With $\hat{\beta}_n$, a sequential design scheme is defined and its consistency and asymptotic normality are obtained under suitable conditions. We then deal with the situation for which the link function may be misspecified and the estimator of the scale parameter may be inconsistent.

Now consider the model with link function (4.1) and variance $V(x - \alpha)$. Let ψ be, as before, the weight function. If β were known, then one would follow (2.2) to define the design sequence $\{x_n, n \geq 1\}$ by

$$\sum_{i=1}^n \psi(\beta x_i)[y_i - H(\beta(x_i - x_{n+1}))] = 0. \quad (4.2)$$

In ignorance of β , we can substitute β in (4.2) by its updated estimator $\hat{\beta}_n \in \mathcal{F}_{n-1}$, the σ -field generated from $y_i, x_i, i \leq n - 1$, to get x_{n+1} :

$$\sum_{i=1}^n \psi(\hat{\beta}_n x_i)[y_i - H(\hat{\beta}_n(x_i - x_{n+1}))] = 0. \quad (4.3)$$

For technical reasons, we also consider an alternative algorithm

$$\sum_{i=1}^n \psi(\hat{\beta}_i x_i)[y_i - H(\hat{\beta}_i(x_i - x_{n+1}))] = 0. \quad (4.4)$$

To analyze the convergence of x_n defined by either (4.3) and (4.4), we need to introduce variants of the conditions (C3) and (C4):

- (C3') $\liminf_{|t| \rightarrow \infty} \frac{|H(\beta t) - p|}{V(t)\psi(t+\alpha)} > 0$ and
 $\liminf_{|t| \rightarrow \infty} \frac{|H(\lambda\beta t) - H(\beta t)|}{V(t)\psi(t+\alpha)} > 0$ for some $1 < \lambda < 2$.
- (C4') V is continuous at 0, H is twice continuously differentiable in a neighborhood of 0 and $H'(0) > 0$.

Theorem 4. *Suppose that conditions (C1), (C2) and (C3') are satisfied and that $\{x_n\}$ is defined either by (4.3) or (4.4).*

- (i) *With probability 1, either $x_n \rightarrow \alpha$, or $x_n \rightarrow \infty$, or $x_n \rightarrow -\infty$,*
 (ii) *If $\liminf_{t \rightarrow \infty} \psi^2(t)[1 + V(t - \alpha)] > 0$, then $P(x_n \rightarrow \infty) = 0$;
 if $\liminf_{t \rightarrow -\infty} \psi^2(t)[1 + V(t - \alpha)] > 0$, then $P(x_n \rightarrow -\infty) = 0$.*
 (iii) *If $\lim_{|t| \rightarrow \infty} \psi^2(t)[1 + V(t - \alpha)] > 0$, then $x_n \rightarrow \alpha$ a.s.*
 (iv) *For the x_n defined by (4.4), if (C4') is also satisfied, $x_n \rightarrow \alpha$ a.s. and $\hat{\beta}_n \in \mathcal{F}_{n-1}$, then*

$$\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N(0, [\beta H'(0)]^{-2} V(0)).$$

Proof. We can apply the same arguments as in the proof of Theorem 1 to show (i) and (ii). The details are therefore omitted. Part (iii) follows directly from (i) and (ii).

To prove (iv), we employ an idea used in Section 3, i.e., by showing that the problem is equivalent to a corresponding stochastic approximation problem. From (4.4), we have

$$\sum_{i=1}^n \psi(x_i) [H(\hat{\beta}_i(x_i - x_{n+1})) - H(\hat{\beta}_i(x_i - x_n))] = \psi(x_n)(y_n - p).$$

Thus

$$x_{n+1} = x_n - \hat{J}_n^{-1}(\psi(x_n)(y_n - p)), \quad (4.5)$$

where $\hat{J}_n(u) = \sum_{i=1}^n \psi(x_i) [H(\hat{\beta}_i(x_i - x_n - u)) - H(\hat{\beta}_i(x_i - x_n))]$. As in the proof of Theorem 2, the first and second derivatives of \hat{J}_n^{-1} are respectively of the orders n^{-1} and n^{-2} . Thus (4.5) can be rewritten as

$$\hat{\alpha}_n = \hat{\alpha}_{n-1} - \frac{1}{nb_n^*} \psi(x_n)(y_n - p)$$

with b_n^* satisfying $b_n^* - n/J'_n(0) = o(n^{-\frac{1}{2} - \epsilon_0})$ for some $\epsilon_0 > 0$. Since $J'_n(0) \in \mathcal{F}_{n-1}$ and $n/J'_n(0) \rightarrow 1/[\beta H'(0)]$, we can again apply Theorem 4 (iv) of Lai and Robbins (1979) to obtain the desired convergence in distribution.

We now relax the restrictions so that (i) H may be misspecified, and (ii) $\hat{\beta}_n$ may not be consistent. We shall show that, by the truncation method as given in (3.1), we can still guarantee the consistency of the design sequence x_n .

To be specific, let $\hat{\beta}_n \in \mathcal{F}_{n-1}$ be a sequence of estimators of β such that for some $\epsilon_0 > 0$, $\hat{\beta}_n \in [\epsilon_0, \epsilon_0^{-1}]$. Define x_n and $\hat{\alpha}_n$ recursively by

$$\sum_{i=1}^n \psi(\hat{\beta}_i x_i) [y_i - \tilde{H}(\hat{\beta}_i(x_i - \hat{\alpha}_n))] = 0,$$

$$x_{n+1} = \max \{ \min(\hat{\alpha}_n, b), a \}, \quad (4.6)$$

where \tilde{H} is twice continuously differentiable, $\tilde{H}'(t) > 0$ for all $t \in [a - \alpha, b - \alpha]$ and $\tilde{H}(0) = p$. Assume that there exists a constant $\tilde{\Delta}$ such that $\tilde{H}(-\tilde{\Delta}) \leq H(\beta(x - \alpha)) \leq \tilde{H}(\tilde{\Delta})$ for all $x \in [a, b]$. Furthermore, $H(0) = p$, H is strictly increasing at 0 and $\sup_{t \in [a, b]} E[|y|^4 | x = t] < \infty$.

Theorem 5. *Suppose that \tilde{H} and H satisfy the preceding assumptions. Let x_n be defined by (4.6).*

(i) $x_n \rightarrow \alpha$ a.s.

(ii) If $\hat{\beta}_n \rightarrow \beta H'(0)/\tilde{H}'(0)$ a.s., then $\sqrt{n}(x_n - \alpha) \xrightarrow{\mathcal{L}} N(0, V(0)/[\beta H'(0)]^2)$.

The proof is omitted since it is similar to those of Theorems 2 and 3.

References

- Blum, J. R. (1954). Approximation methods which converge with probability one. *Ann. Math. Statist.* **25**, 382-386.
- Chow, Y. S. (1965). Local convergence of martingales and the law of large numbers. *Ann. Math. Statist.* **36**, 552-558.
- Chung, K. L. (1954). On a stochastic approximation method. *Ann. Math. Statist.* **25**, 463-483.
- Finney, D. J. (1978). *Statistical Method in Biological Assay*. Griffin, London.
- Ford, I., Torsney, B. and Wu, C. F. J. (1992). The use of a canonical form in the construction of locally optimal designs for nonlinear problems. *J. Roy. Statist. Soc. Ser. B* **54**, 569-583.
- Goodwin, G. C., Ramadge, P. J. and Caines, P. E. (1981). Discrete time stochastic adaptive control. *SIAM J. Control Optim.* **19**, 829-853.
- Kumar, P. R. (1985). A survey of some results in stochastic adaptive control. *SIAM J. Control Optim.* **23**, 329-380.
- Lai, T. L. and Robbins, H. (1979). Adaptive design and stochastic approximation. *Ann. Statist.* **7**, 1196-1221.
- Lord, F. M. (1971). Tailored testing, an application of stochastic approximation. *J. Amer. Statist. Assoc.* **66**, 707-711.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. (2nd Ed.) Chapman and Hall, London, New York.
- Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer-Verlag, New York.
- Robbins, H. and Monro, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22**, 400-407.
- Robbins, H. and Siegmund, D. (1971). A convergence theorem for non-negative almost supermartingales and some applications. In *Optimizing Methods in Statistics* (Edited by J. Rustagi), 233-257. Academic Press, New York.
- Sacks, J. (1958). Asymptotic distribution of stochastic approximation procedures. *Ann. Math. Statist.* **29**, 373-405.

- Sitter, R. R. and Torsney, B. (1995). D-Optimal designs for generalized linear models. In *MODA4-Advances in Model Oriented Data Analysis* (Edited by C. P. Kitsos and W. G. Muller), 87-102. Physica-Verlag Heidelberg.
- Wu, C. F. J. (1985). Efficient sequential designs with binary data. *J. Amer. Statist. Assoc.* **80**, 974-984.
- Wu, C. F. J. (1986). Maximum likelihood recursion and stochastic approximation in sequential designs. In *Adaptive Statistical Procedures and Related Topics* Vol. 8 (Edited by J. Van Ryzin), 298-314. IMS Monograph Series.
- Wu, C. F. J. (1988). Optimal design for percentile estimation of a quantal response curve. In *Optimal Design and Analysis of Experiments* (Edited by Y. Dodge, V. Federov and H. P. Wynn), 213-224. Elsevier Science Publishers B. V., (North Holland).

Department of Statistics, Hill Center, Busch Campus, Rutgers University, New Brunswick, NJ 08903, U.S.A.

Department of Statistics, Mason Hall, University of Michigan, Ann Arbor, MI 48109-1027, U.S.A.

(Received August 1995; accepted February 1996)