MRC Technical Summary Report #2692

EFFICIENT SEQUENTIAL DESIGNS WITH BINARY DATA

C. F. Jeff Wu

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

May 1984

(Received May 2, 1984)

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

EFFICIENT SEQUENTIAL DESIGNS WITH BINARY DATA

C. F. JEFF WU*

Technical Summary Report #2692

May 1984

ABSTRACT

A class of sequential designs for estimating the percentiles of a quantal response curve is proposed. Its updating rule is based on an efficient summary of all the data available via a parametric model. The "logit-MLE" version of the proposed designs can be viewed as a natural analogue of the Robbins-Monro procedure in the case of binary data. It is shown to be asymptotically consistent, distribution-free and optimal via its connection with the latter procedure. For certain choices of initial designs the proposed method performs very well in a simulation study for sample sizes up to 35. A nonparametric sequential design, via the Spearman-Kärber estimator, for estimating the median is also proposed.

AMS (MOS) Subject Classifications: 62K05, 62L05

Key Words: Logit, Optimal design, Quantal response curve, Robbins-Monro stochastic approximation, Sensitivity experiments, Spearman-Kärber estimator, Up-and-Down method

Work Unit Number 4 - Statistics and Probability

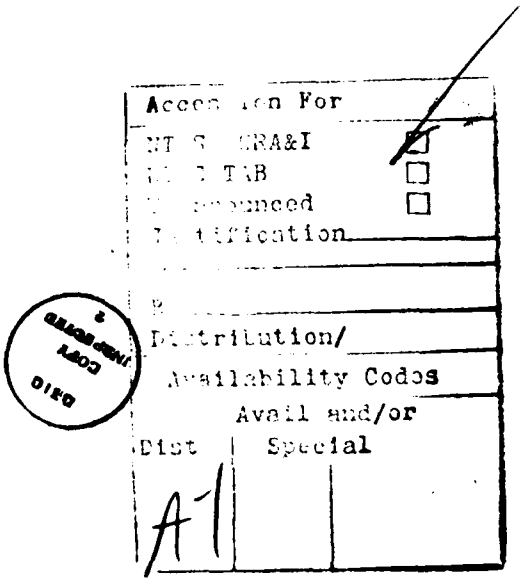
---

* Work completed while visiting the Mathematical Sciences Research Institute, Berkeley.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and the ARO Grant No. DAAG29-82-K0154.

SIGNIFICANCE AND EXPLANATION

In many physical or biological experiments with binary response a quantal response curve is assumed to relate the probability of response to the corresponding level of the stimulus variable. To estimate the percentiles of the quantal response curve efficiently, a sequential design is often used in practice. We propose a new class of sequential designs with updating rules based on an efficient summary of all data available via a parametric model. This method is shown to be asymptotically as good as the optimal stochastic approximation method. More importantly, its finite sample performance in a simulation study is often better than the latter method. For fixed initial designs, the percentage of runs saved by using our method ranges from 25% to 57%.



EFFICIENT SEQUENTIAL DESIGNS WITH BINARY DATA

C. F. Jeff Wu*

1. Introduction

A sensitivity experiment is characterized by a response curve that relates the stimulus level applied to an experimental subject to the probability of response. The outcome of the experiment is assumed dichotomous, response or nonresponse. This situation arises in many fields of research. In testing the strength of materials, the stimulus level may be the level of impact energy applied to a piece of material, and the response is either "fail" or "not fail" (Wetherill, 1963). In testing explosives, the stimulus level may be the height from which a weight is dropped or the pressure directly applied to the explosive, and the response is "explode" or "not explode" (Dixon and Mood, 1948). In biological assays a test animal survives or not at a given dose level (Finney, 1978). In psycho-physical research the probability of detecting a stimulus is related to its intensity level (Rose et al., 1970). In educational testing, one may want to study the "item characteristic curve" that relates the difficulty level of the test item to the probability of "right" or "wrong" answer (Lord, 1971).

Our main interest is in estimating the percentiles of the response curve $F(x)$, which is the probability of response for a given stimulus level x. The 100p percentile L_p is defined as

$$F(L_p) = p \quad . \tag{1}$$

---

* Work completed while visiting the Mathematical Sciences Research Institute, Berkeley.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and ARO Grant No. DAAG29-82-K0154.

For simplicity we assume F is monotone increasing and continuous. The median of F, $L_{0.5}$, is the most commonly used measure of a characteristic of the response curve. In some situations estimating $L_{0.5}$ is of intrinsic interest, but more often it is because $L_{0.5}$ is easy to estimate. In quality assurance it may be more relevant to study the extreme percentiles, e.g., to find the impact energy level that results in the failure of material for at most 10% of the time. On the other hand $L_{0.9}$ may be more relevant in explosive research.

In this paper we will present some new sequential designs for the efficient estimation of L_p for small or moderate sized experiments. The sequential designs are constructed in such a way that all the information in the previous runs can be efficiently utilized in suggesting how the next run should be conducted. When the experimental runs are very expensive, the saving of a few runs by an efficient design outweighs the extra pains taken in designing a sequential experiment. The sequential nature of the design requires quick responses so that the experiment will not be unduly prolonged. It is suitable, for example, when the experimental facility is limited so that the experiments must be performed one after another. It is not applicable to many biological experiments that involve inexpensive animals and slow responses. Therefore our method is more appropriate for expensive experiments with short response time, which are more often encountered in engineering research. In educational or psychological testings, if a test has to be repeated routinely on many subjects, it pays off to automate the design and to look for the most efficient ones (in terms of reducing the number of test items).

In the next section we shall review two nonparametric sequential designs (the Robbins-Monro and the Up-and-Down methods) with special reference to small sample binary experiments. Our approach is to assume a parametric model

for the response curve and estimate efficiently the relevant parameters in the model based on all the data available. An estimated quantal response curve (EQRC) is constructed through the current estimate of the parameters and the next design point is determined from the EQRC. If the two-parameter logistic model is used and the parameters are estimated by the maximum likelihood, we call it a "logit-MLE" design. It is demonstrated heuristically to be a natural analogue in the case of binary data, of the Robbins-Monro (RM) procedure for continuous data. Its consistency is proved under two sets of restrictive conditions. Assuming consistency, it is shown to be asymptotically equivalent to an adaptive Robbins-Monro procedure. Therefore it is asymptotically distribution-free and optimal (in a sense defined in Section 2), two properties enjoyed by the latter procedure. Truncated versions, (10) and (21), of the two methods are considered for the purpose of stabilizing their performance. They are compared with the nonadaptive RM procedure in a simulation study for sample sizes up to 35. If a fixed initial design with wide-spread design levels is chosen, the logit-MLE design seems to take full advantage of the information in the past data. It substantially outperforms the adaptive RM procedure, which in turn outperforms the nonadaptive RM procedure. On the other hand, if a (nonadaptive) RM procedure is used in generating the initial design levels, the relative performance of the nonadaptive RM design and the two adaptive designs (RM and logit-MLE) depends on the starting value x_1 and the constant c (formula (2)) in the RM procedure. For good choices of x_1 and c (specified in Section 7), which usually requires some prior knowledge about the unknown response curve, the nonadaptive RM design performs very well. In the absence of such knowledge, the two adaptive designs perform better. Detailed comparisons and further remarks are found in Sections 6 and 7. A nonparametric sequential design for

estimating the median $L_{0.5}$ is proposed via the Spearman-Kärber estimator. Its limitations are discussed.

2. Review and criticism of the Stochastic Approximation method and the Up-and-Down method.

The Stochastic Approximation method and the Up-and-Down method are two most commonly used nonparametric sequential designs for quantal response problems.

Stochastic Approximation Method (Robbins and Monro, 1951):

Let $y_n = 1$ or 0 as the n^{th} experiment results in a response or non-response. For estimating L_p, the stimulus level x_{n+1} of the $(n+1)^{th}$ run is chosen according to

$$x_{n+1} = x_n - \frac{c}{n}(y_n - p) \quad . \tag{2}$$

According to the results of Chung (1954), Hodges and Lehmann (1955) and Sacks (1958), it is optimal to choose c in (2) to be equal to $(F'(L_p))^{-1}$ in the sense of minimizing the asymptotic variance of $\sqrt{n}(x_n - L_p)$ within the class (2). Except for normal errors, the resulting procedure is not asymptotically fully efficient, that is, its asymptotic variance does not achieve the Cramer-Rao lower bound. Abdelhamid (1973) and Anbar (1973) proposed to transform $y_n - p$ in order that the asymptotic variance of $\sqrt{n}(x_n - p)$ be minimized. However the optimal transformation depends on the distribution of the errors, which is typically unavailable to the experimenter in the situations under consideration. For the rest of the paper, any procedure that achieves minimal asymptotic variance within the class (2) will be called asymptotically optimal without special reference to (2).

The small sample behavior of (2) depends very much on a good starting value x_1 (Wetherill, 1963). Ideally x_1 should be close to L_p. A good guess of the optimal constant c may also be hard to come by since in most

practical situations the experimenters have little idea about the slope of F at L_p. Poor choice of c and x_1 will make (2) an inefficient procedure for small and even moderate samples. The stochastic approximation method has been used more effectively in on-line estimation wherein a large number of data have to be processed quickly.

To achieve minimal asymptotic variance within the class (2), it is necessary to estimate the slope $F'(L_p)$. One such estimator is the regression slope of y_i over x_i,

$$\hat{\beta}_n = \frac{\sum_1^n y_i(x_i - \bar{x}_n)}{\sum_1^n (x_i - \bar{x}_n)^2}, \quad \bar{x}_n = \frac{1}{n}\sum_1^n x_i \ ,$$

which gives the following adaptive Robbins-Monro procedure

$$x_{n+1} = x_n - \frac{1}{n\hat{\beta}_n}(y_n - p) \ . \tag{2a}$$

Under various regularity conditions, Anbar (1978) and Lai and Robbins (1981) proved that $\hat{\beta}_n$ converges to $F'(L_p)$ and the procedure (2a) has the same asymptotic distribution as the optimal nonadaptive Robbins-Monro (RM) procedure (2) with $c = (F'(L_p))^{-1}$.

The RM procedure can be given a finite sample justification if y and x are related via a simple linear regression model

$$y_i = \alpha + \beta x_i + e_i$$

where e_i are i.i.d. normally distributed with mean zero and variance σ^2. Assume β is known and the parameter of interest is $\theta = -\alpha/\beta$, the solution of the linear equation $\alpha + \beta x$. Then it can be shown that (Lai and Robbins, 1979, Lemma 1) the RM procedure

$$x_{n+1} = x_n - \frac{c}{n} y_n \ , \quad c = \beta^{-1} \tag{3}$$

is equivalent to the procedure

$$x_{n+1} = \bar{x}_n - \beta^{-1}\bar{y}_n = -\hat{\alpha}_n/\beta \ , \tag{3}'$$

where $\hat{\alpha}_n = \bar{y}_n - \beta\bar{x}_n$ is the <u>maximum likelihood estimate</u> (MLE) of α. There-

fore the next design point x_{n+1} according to (3)' is the solution of the estimated linear equation

$$\hat{F}_n(x) = \hat{\alpha}_n + \beta x, \ \hat{\alpha}_n = \text{MLE of } \alpha .$$

Although this justification is specific to the linearity of Ey in x and the normality of error e_i, the consistency and asymptotic normality of the non-adaptive RM procedure (3) hold under much weaker conditions. The equivalence of (3) and (3)' breaks down when β is replaced by $\hat{\beta}_n$. Since $\hat{\beta}_n$ is close to β for large n, it may not be unreasonable to interpret the adaptive RM procedure (2a) with $p = 0$ as an approximation to the solution of the estimated linear equation

$$\hat{\alpha}_n + \hat{\beta}_n x, \ (\hat{\alpha}_n, \hat{\beta}_n) = \text{MLE of } (\alpha, \beta)$$

for the simple linear regression model and large n. For the problem of estimating $L_p = F^{-1}(p)$, it can be viewed as a stochastic version of the Newton-Raphson method for solving $F(x) = p$ by the tangential approximation to F at $\bar{x}_n$ with $F(\bar{x}_n)$ replaced by $\bar{y}_n$ and $F'(\bar{x}_n)$ by $\hat{\beta}_n$. The procedure (2a) is aptly called a "stochastic Newton-Raphson" method in Anbar (1978).

When n is small or the current guess x_n is on the tails of the response curve, $\hat{\beta}_n^{-1}$ may behave erratically. Since the tails of the response curve are flat, $\hat{\beta}_n$ with $\{x_i\}_1^n$ located on either tail tends to be closer to zero, thus making the adjustment from x_n to x_{n+1} in (2a) unreasonably large. (It reminds us of the well-known numerical instability of the Newton-Raphson method for solving the quantal response equation $F(x) = p$ when the starting value x_1 is on the tails of F.) This happens when the initial guess is poor for estimating the median or when the initial design takes too few points on the middle part of the response curve for estimating the extreme percentiles. To remedy this, we propose to truncate $\hat{\beta}_n^{-1}$, that is, to use $\max(\min(\hat{\beta}_n^{-1}, d), \delta)$ instead of $\hat{\beta}_n^{-1}$ for some positive constants $d > \delta$. The

simulation study of Section 6 shows that there is considerable improvement in using this truncated version of (2a).

Up-and-Down method (Dixon and Mood, 1948):

$$x_{n+1} = \begin{cases} x_n + \Delta \\ x_n - \Delta \end{cases} \quad \text{if} \quad y_n = \begin{cases} 0 \\ 1 \end{cases} . \tag{4}$$

The method works only for $L_{0,5}$. It is very simple to implement but, for small or moderate samples, its performance depends very much on a good guess of x_1 and Δ. Unless the step size Δ is made adaptive, the large sample property of x_n cannot be studied. Its empirical performance is usually not as good as the Stochastic Approximation method. This and several modifications of the two methods can be found in Wetherill (1963, 1966).

Both methods are "Markovian" in that the choice of the next run depends sensibly on the outcome of the current one. Their simplicity was a crucial factor when inexpensive computing was not accessible. Their main disadvantages are: (a) The updating rules (2) and (4) do not make use of all the data available in an efficient way, and thus making the choice of step size less flexible. (b) Their small sample behavior depends on a good choice of the relevant constants in (2) and (4), which in turn depends on the experimenter's knowledge of the unknown response curve F. For small or moderate sized experiments with expensive runs, inefficiency and lack of robustness can be quite serious. Large sample properties, which depend on locally linear approximations, are not always relevant in this context.

To overcome the shortcomings (a) and (b), an alternative method is proposed in the next section.

3. A class of sequential designs based on the estimated quantal response curve.

Ideally we would like to have a good estimate $\hat{F}_n$ of the whole curve F, from which the next design point x_{n+1} is chosen to be its $100p$ percentile, i.e., $\hat{F}_n(x_{n+1}) = p$. A (smooth) nonparametric estimate $\hat{F}_n$ of F is not feasible since it requires a large number of observations for $\hat{F}_n$ to be a good estimate. A natural approach for small sample problems is to assume a parametric model

$$F(x) = H(x|\theta),\ H \text{ is continuous in } x\ ,$$

$$\lim_{x\to-\infty} H(x|\theta) = 0,\ \lim_{x\to\infty} H(x|\theta) = 1\ .$$

The general recipe of our sequential design procedure for estimating L_p is:

(i) find an efficient estimate $\hat{\theta}_n = \hat{\theta}((y_i,x_i)_1^n)$ of θ ,

(5)

(ii) define the estimated quantal response curve (EQRC)

$$\hat{F}_n(x) = H(x|\hat{\theta}_n)\ ,$$

and choose the next design x_{n+1} s.t. $\hat{F}_n(x_{n+1}) = p$.

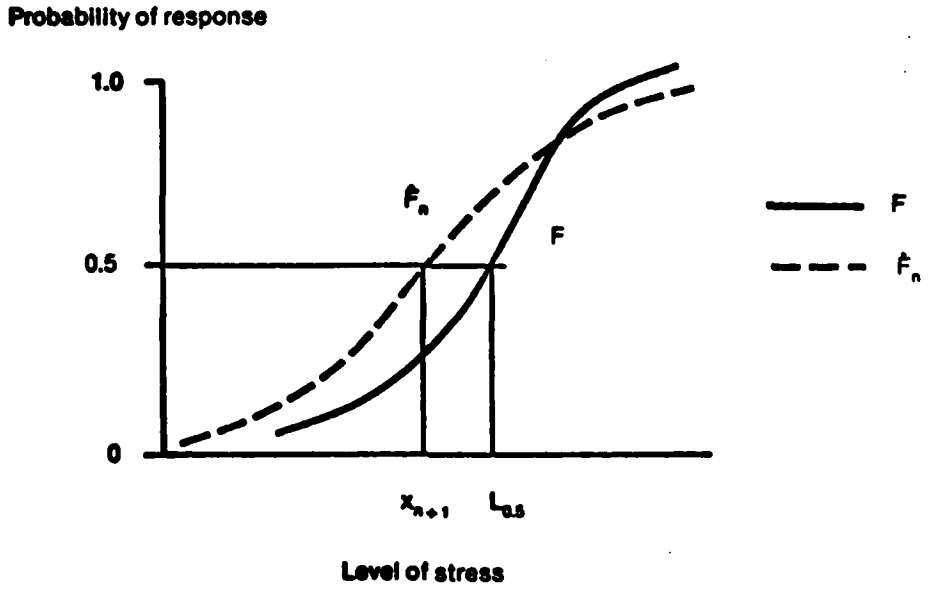


Figure 1. A representation of procedure (5) with p = 0.5

If $H(x|\theta)$ in (5) equals $\alpha + \beta x$, β known, and the continuous measurement y is related to x through a simple linear regression model with normal error, the procedure (5) with the maximum likelihood estimate is identical to the nonadaptive RM procedure (2) with $c = \beta^{-1}$. See the discussion following (3) and (3)'. In this sense our proposal can be viewed as a natural analogue of the RM procedure for binary data. Since the straight line model $Ey = \alpha + \beta x$ provides a finite sample justification for the RM method for <u>continuous</u> y, it would seem natural to use the two-parameter logit curve

$$H(x|\theta) = (1 + e^{-\lambda(x-\alpha)})^{-1}, \quad \lambda > 0, \ \theta = (\alpha,\lambda) \tag{6}$$

for modelling the <u>binary</u> response y and the stimulus level x in procedure (5). If λ in (6) is chosen to be a known constant, the resulting procedure (5) is nonadaptive. When α and λ are both estimated from the data, (5) is adaptive.

If the experimenter has some knowledge about his problem, it should be taken into account in the choice of the parametric model $H(\ |\theta)$. Given this model were there a reliable prior on θ, a Bayesian approach (Freeman, 1970; Tsutakawa, 1972; Owen, 1975; Leonard, 1982) for estimating θ would be appropriate. In the absence of such information, it seems appropriate to use a simple model like the logit or probit.

The main reason for preferring logit to its competitor, the probit model,

$$G(x|\theta) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-\frac{(z-\mu)^2}{2\sigma^2}}\, dz, \quad \theta = (\mu,\sigma), \ \sigma > 0 \ , \tag{7}$$

is computational ease. It is well known that the logit, the probit and other parametric models like the angular and the linear curves agree very closely in the range 0.2 to 0.8 (Cox, 1970, Table 2.1). We do not see any advantage in using the probit over the logit. It is rarely the case that a parametric quantal response model be justifiable on biological or physical grounds. The

successful use in practice of the parametric approach for quantal response problems is mainly due to this key fact that the parametric curves (after adjusting for location and scale) agree very closely in a wide range of p values.

For p outside $[0.1, 0.9]$ the percentiles for different parametric models vary greatly. The choice between (5) and (2) (or (2a)) is not clear-cut. The procedure (5) is more vulnerable to the misspecification of F. On the other hand, the RM procedure (2) or (2a) will require a very large sample size to become distribution-free. For instance, the method makes on the average nine negative moves for each positive move in the neighborhood of $L_{0.9}$. Instead of "straddling" $L_{0.9}$, the sequence makes far too many moves in one direction. This may explain the much poorer empirical performance of the RM method (2) even for moderate percentiles like $L_{0.75}$ (Wetherill, 1963, §6).

The next issue is the choice of efficient estimator $\hat{\theta}_n$ in (5) (i). The minimum logit chi-square method (Berkson, 1955) is not suitable for the kind of data generated by a sequential procedure like (5), especially for small or moderate samples. This is because there are few, and typically only one or two, observations at a given x level to make the minimum logit chi-square work. Unless we restrict the search of design levels to a small number of x levels, the situation will not change much. The same remark applies to the minimum modified chi-square method, and to a lesser extent, to the minimum chi-square method. The maximum likelihood estimate of (α,λ) in (6) is obtained by iteratively solving the equations

$$\text{(8)} \qquad \begin{aligned} \sum_{i=1}^{n} H(x_i|\alpha,\lambda) &= \sum_{1}^{n} y_i \ , \\ \sum_{1}^{n} x_i H(x_i|\alpha,\lambda) &= \sum_{1}^{n} y_i x_i \ , \end{aligned}$$

where $H(x|\alpha,\lambda) = (1 + e^{-\lambda(x-\alpha)})^{-1}$. The MLE $(\hat{\alpha},\hat{\lambda})$ is a function of the sufficient statistics $(\Sigma\, y_i, \Sigma\, y_i x_i)$ and is asymptotically efficient given the right model. Under (6),

$$L_p = \alpha - \frac{1}{\lambda} \ell n(\frac{1}{p} - 1)$$

$$\hat{L}_p = \hat{\alpha} - \frac{1}{\hat{\lambda}} \ell n(\frac{1}{p} - 1) \quad .$$

For the implementation of (5), it is important to know when the MLE exists. Assume there are at least two distinct x_i's. It is known (Silvapulle, 1981) that the MLE of the "linear" parameters $(\lambda,\lambda\alpha)$ in the logit model (6) exists uniquely if and only if the following "interlocking" condition is satisfied,

(9.1) $\quad (x^+_{min}, x^+_{max}) \cap (x^-_{min}, x^-_{max})$ is non-empty

or

(9.2) $$x^+_{min} < x^-_{min} = x^-_{max} < x^+_{max}$$

or

(9.3) $$x^-_{min} < x^+_{min} = x^+_{max} < x^-_{max}$$

where $x^+_{max(min)} = \max(\min)\ \{x_i : y_i = 1\}$, $x^-_{max(min)} = \max(\min)\ \{x_i : y_i = 0\}$. The same result holds for more general distributions F including the probit model (7). See Silvapulle (1981, Theorem (iii)). It is easy to see that (9), once satisfied, is always satisfied by the addition of more observations.

If the MLE is chosen for (5 i), it is critical not to start the iteration in (5) until the condition for the existence and uniqueness of the MLE is satisfied. A <u>premature start</u> of the procedure (5) will lead to <u>inconsistent estimate</u> as the following example shows. When (9.1) - (9.3) are violated, the two intervals $[x^-_{min}, x^-_{max}]$ and $[x^+_{min}, x^+_{max}]$ separate or share one point in common. <u>Any</u> point in $[x^-_{max}, x^+_{min}]$ (or $[x^+_{max}, x^-_{min}]$ whichever applies) maximizes the likelihood. Take an observation at any such point will again

violate (9.1) - (9.3). By repeating (5), a sequence of x_n will be obtained that lie in the initial interval $[x^-_{max}, x^+_{min}]$. If the initial interval does not contain the true parameter L_p, the estimate x_n for any large n will never be close to L_p.

The question of conducting the initial runs before there is a unique MLE, is very difficult unless some prior knowledge is available. It may be done in an ad hoc manner aided with experience, or by the RM procedure (2) with a reasonable guess of x_1 and a slightly larger c than the experimenter's guess. Wetherill (1963) showed that the procedure (2) with larger c is less susceptible to a poor choice of x_1 especially for small samples (see also the discussion of Table IV).

The change from x_n to x_{n+1} via the logit-MLE method may be unduly large when the problem is "ill-posed." It happens in the first few runs after the existence and uniqueness of the MLE is first satisfied. We propose a truncated version as follows. Define d_n as the solution of $x_{n+1} = x_n - \frac{d_n}{n}(y_n - p)$, where $x_{n+1} = \hat{\alpha}_n - \hat{\lambda}_n^{-1} \ell n(p^{-1}-1)$ and $(\hat{\alpha}_n, \hat{\lambda}_n)$ is the solution of (8). The $(n+1)^{th}$ design level is chosen to be

$$x_n - \frac{d_n^*}{n}(y_n - p), \quad d_n^* = \max(\delta, \min(d_n, d)), \quad d > \delta > 0 \quad . \tag{10}$$

According to the simulation results, this truncation turns out to be very effective.

Since the logit (and any other parametric) assumption is vulnerable on the extreme tails, it may be desirable to use an estimation method that places less weight on the observation with more extreme x_i. For data generated by sequential procedures like (2), (4) and (5), the x_i's in the initial runs tend to be more extreme. A simple way to achieve this is to insert weight $w_i = w(|x_i - x_n|)$ on both sides of (8) and solve iteratively the weighted version of the likelihood equation (8), where $w(z)$ is decreasing in $z > 0$,

and x_n is considered to be a good estimate of L_p. If we choose w_i to be 0 or 1, it is equivalent to performing the unweighted MLE based on a subset of data with moderate x_i's. The general question of robut estimation for quantal response data was addressed in Miller and Halpern (1980).

For small n we advocate the use of a simple model like the logit for procedure (5) since it is difficult to discriminate between two binary response models (Chambers and Cox, 1967). For larger n, a symmetric logit or probit model will not be appropriate if the true F is skewed. A three-parameter model may be used in (5) when the data indicate that the additional skewness parameter is indeed significant. This will make the procedure (5) less susceptible to the incorrect initial choice of the parametric model. A skewed logit model, (22), will be considered in Section 6.

An important question, which is beyond the scope of the paper, concerns the time to terminate the experiment with adequate information. Let $\hat{v}$ be an estimate of the variance $\text{var}(\hat{L}_p^{(n)})$ via the assumed parametric model, where $\hat{L}_p^{(n)}$ is the MLE of L_p from the first n observations. A stopping rule may be devised based on the value of $\hat{v}$.

4. A sequential design for estimating $L_{0.5}$ based on the Spearman-Kärber estimator

If the unknown response curve $F(x) = H(x-\alpha,\phi)$ is skew-symmetric about α, i.e. $H(z,\phi) + H(-z,\phi) = 2H(0,\phi)$ for any z, ϕ, α is both the median $L_{0.5}$ and the mean of F. The Spearman-Kärber estimator (Finney, 1978, p. 394) is a nonparametric estimator of the (discretized) mean of F,

$$\hat{\alpha}_{SK} = \sum_{j=1}^{J} (\hat{p}_j - \hat{p}_{j-1}) \frac{1}{2} (x_{j-1} + x_j) \quad ,$$

where $x_1 < \ldots < x_J$, n_j observations are taken at x_j with r_j responses, $\hat{p}_j = r_j/n_j$, $n = \Sigma_1^J n_j$. Under conditions that ensure that $\hat{\alpha}_{SK}$ is an

efficient estimator of α, an alternative sequential design for estimating the median $L_{0.5} = \alpha$ is the following:

$$\text{(i) compute } \hat{\alpha}_{SK}^{(n)} = \hat{\alpha}_{SK}((y_i, x_i)_1^n) \ , \tag{11}$$

$$\text{(ii) set } x_{n+1} = \hat{\alpha}_{SK}^{(n)} \ .$$

The two distinct advantages of the procedure (11) are: 1) computational ease, 2) weak assumption on F, i.e., the functional form of H is not assumed known. But the price to pay for these is quite dear. The conditions required to ensure a proper performance of (11) are quite restrictive. First, F should be skew-symmetric so that its mean and median are equal. Since $\hat{\alpha}_{SK}$ is an unbiased estimator of the discretized mean, not the population mean, their difference becomes negligible only when the spacing $\{x_i\}_1^n$ is reasonably dense. A proper use of $\hat{\alpha}_{SK}$ requires that x_1 and x_J are chosen such that $F(x_1) = 0$, $F(x_J) = 1$, which may be hard to achieve in the initial stage of the type of sequential designs considered in the paper. If the experimenter has to pray for the validity of these assumptions, the procedure (11) can not be truly "nonparametric." Therefore it will not be included in the empirical study.

5. Some large sample results concerning the logit-MLE version of the design (5)

In this section some theoretical properties of the "logit-MLE" version of (5) are investigated. Two consistency results are established under rather restrictive conditions. Assuming consistency, it will be shown that the "logit-MLE" version of (5) is asymptotically equivalent to the adaptive Robbins-Monro procedure (2a). Since the latter is nonparametric and is asymptotically optimal within the class of methods in (2), the former is optimal in the same sense whether the true F function is logistic or not. For those

who wonder why a model-based procedure turns out to be asymptotically distribution-free, we merely recall the fact that the Robbins-Monro procedure can be formally viewed as a special case of (5) with $F(x)$ being a linear function in x, although linearity does not play a role in the asymptotic behavior of the RM procedure.

First we establish the equivalence of the nonadaptive "logit-MLE" version of (5) and the nonadaptive RM procedure for estimating the median $L_{0.5}$. Without loss of generality, we assume the scale parameter λ in (6) equals 1. According to (5), we have to solve the first equation of (8) for choosing x_n and x_{n+1}, i.e.,

$$\sum_{i=1}^{n-1} \frac{1}{1+e^{-(x_i-x_n)}} = \sum_{1}^{n-1} y_i \tag{12.1}$$

$$\sum_{i=1}^{n} \frac{1}{1+e^{-(x_i-x_{n+1})}} = \sum_{1}^{n} y_i \quad . \tag{12.2}$$

By substracting (12.1) from (12.2) and after some algebras, we obtain

$$\sum_{i=1}^{n} \frac{e^{x_n-x_i}(1-e^{x_{n+1}-x_n})}{(1+e^{x_n-x_i})(1+e^{x_{n+1}-x_n}e^{x_n-x_i})} = y_n - \frac{1}{2} \quad , \tag{13}$$

which defines $x_{n+1} - x_n$ implicitly as a function of $y_n - \frac{1}{2}$ and x_i, $i = 1,\ldots,n$. Let $A(w)$ denote the left hand expression of (13) as a function of $w = x_{n+1} - x_n$. It is easily verified that

$$0 < -A'(w) = \sum_{1}^{n} \frac{e^{x_n-x_i+w}}{\left(1+e^{x_n-x_i+w}\right)^2} < \frac{n}{4} \quad .$$

Since $A(w)$ is monotone decreasing, $x_{n+1} - x_n < 0$ for $y_n = 1$ and > 0 for $y_n = 0$. Since $A(0) = 0$, (13) can be rewritten as

$$\int_{x_{n+1}-x_n}^{0} (-A'(w))dw = y_n - \frac{1}{2} \quad \text{for} \quad y_n = 1$$

and

$$\int_{0}^{x_{n+1}-x_n} A'(w)dw = y_n - \frac{1}{2} \quad \text{for} \quad y_n = 0 \quad ,$$

which implies, using $-A'(w) \leq n/4$,

$$y_n - \frac{1}{2} \leq -\frac{n}{4}(x_{n+1}-x_n) \quad \text{for} \quad y_n = 1$$

(14)

$$-(y_n - \frac{1}{2}) \leq \frac{n}{4}(x_{n+1}-x_n) \quad \text{for} \quad y_n = 0 \quad .$$

We can now express $x_{n+1} - x_n$ as a Robbins-Monro recursion

$$x_{n+1} = x_n - \frac{b_n}{n}(y_n - \frac{1}{2}) \quad ,$$

where b_n is <u>implicitly</u> defined via the equation $A(w) = \pm \frac{1}{2}$. From (14), $b_n > 4$. By further bounding b_n from above, and modifying the nonadaptive "logit-MLE" version of (5) as follows,

(15) $$x_{n+1} = x_n - \frac{B_n}{n}(y_n - \frac{1}{2}), \; B_n = \min(b_n, B) \quad ,$$

where B is a constant > 4, it follows from standard results (Robbins and Siegmund, 1971) on the consistency of the RM-type recursion that the modified "logit-MLE" design (15) converges to $L_{0.5}$ with probability 1.

A similar modification of the adaptive "logit-MLE" design was considered in (10). We are not able to give a rigorous proof of its consistency, although the simulation results of Section 6 suggest that it should be so. We can prove consistency under the very restrictive condition that the MLE $(\hat{\alpha}_n, \hat{\lambda}_n)$ in (8) converges uniformly to a constant (α^*, λ^*), $\lambda^* \neq 0$, and therefore $x_{n+1} = \hat{\alpha}_n - \hat{\lambda}_n^{-1} \ell n(p^{-1}-1)$ converges uniformly to a constant x^*. More precisely we will prove $x^* = L_p$, that is, x_n converges to the true parameter L_p whether the true F is logit or not. From (5 ii), we have $H(x_{n+1}|\hat{\alpha}_n, \hat{\lambda}_n) = p$, where H is the logit function (6), and as $n \to \infty$ we obtain $H(x^*|\alpha^*, \lambda^*) = p$. On the other hand, the first equation of (8) is equivalent to

(16) $$\frac{1}{n}\sum_{1}^{n} H(x_i|\hat{\alpha}_n,\hat{\lambda}_n) = \frac{1}{n}\sum_{1}^{n} y_i \quad ,$$

whose left hand expression converges to $H(x^*|\alpha^*,\lambda^*) = p$ from the uniform convergence of $\hat{\alpha}_n$ and $\hat{\lambda}_n$. For the convergence of the right hand expression of (16), note that y_i has a binomial distribution with parameter $F(x_i)$, where x_i is measurable with respect to the past $i-1$ observations. From a strong law of large numbers in Dubins and Freedman (1965),

$$\sum_{1}^{n} y_i / \sum_{1}^{n} F(x_i) \to 1 \quad \text{a.s.}$$

which implies $n^{-1}\sum_{1}^{n} y_i \to F(x^*)$ since $n^{-1}\sum_{1}^{n} F(x_i) \to F(x^*)$. By equating the limits of the two sides of (16), we obtain $p = F(x^*)$, or equivalently, $x^* = L_p$.

Assuming the consistency of the "logit-MLE" design sequence x_n, we will prove that it is asymptotically equivalent to the adaptive RM procedure (2a). Consider the approximation

(17) $$J(t) = \frac{1}{1+e^{-t}} \approx p + (t-J^{-1}(p))J'(J^{-1}(p)),\ J^{-1}(p) = -\ell n(\frac{1}{p} - 1) \quad ,$$

which is approximately valid for x_i close to L_p. By applying (17) to (8) we obtain

(18)
$$\sum_{1}^{n} (\lambda x_i - \lambda L_p) = \frac{1}{J'(J^{-1}(p))}\sum_{1}^{n} (y_i - p)$$
$$\sum_{1}^{n} (\lambda x_i^2 - \lambda L_p x_i) = \frac{1}{J'(J^{-1}(p))}\sum_{1}^{n} (y_i - p)x_i$$

where the $100p$ percentile $L_p = \alpha - \frac{1}{\lambda}\ell n(\frac{1}{p} - 1) = \alpha + \frac{1}{\lambda}J^{-1}(p)$. The estimator $\hat{L}_p^{(n)}$, (19), is obtained from $\widehat{\lambda L}_p$ and $\hat{\lambda}$ by solving (18),

(19) $$\hat{L}_p^{(n)} = \frac{\widehat{\lambda L}_p}{\hat{\lambda}} = \frac{\sum_{1}^{n} x_i \sum_{1}^{n} (y_i-p)x_i - \sum_{1}^{n} x_i^2 \sum_{1}^{n} (y_i-p)}{\sum_{1}^{n} 1 \sum_{1}^{n} (y_i-p)x_i - \sum_{1}^{n} x_i \sum_{1}^{n} (y_i-p)} \quad ,$$

which is the weighted average of x_i with weight w_i proportional to $\sum_j^n (y_j-p)(x_j-x_i)$. Since $\hat{L}_p^{(n)}$ is independent of $J'(J^{-1}(p))$, $J'(J^{-1}(p))$ in the approximation (17) can be replaced by any other constant without affecting the subsequent results (19) and (20).

Note that some w_i may be negative. The denominator of (19) is equal to $n \sum_{y_i=1} x_i - \sum_{y_i=1} 1 \sum^n x_i$, which is nonzero unless $\left(\sum_{y_i=1} 1\right)^{-1} \sum_{y_i=1} x_i = n^{-1} \sum^n x_i$. From (19) and after some algebras, it is easy to show that the $(n+1)^{th}$ run, according to the procedure (5),

$$x_{n+1} = \hat{L}_p^{(n)} = \hat{L}_p^{(n-1)} - \frac{(y_n-p)\sum^n (x_i-x_n)^2}{n\sum y_i(x_i-\bar{x}_n)}$$

(20)

$$= x_n - \frac{c_n}{n}(y_n-p),\ c_n = \frac{\sum^n (x_i-x_n)^2}{\sum y_i(x_i-\bar{x}_n)},$$

where $\bar{x}_n = n^{-1} \sum^n x_i$. Therefore the linear approximation (20) to our procedure (5) is asymptotically optimal if c_n in (20) converges almost surely to $[F'(L_p)]^{-1}$. To this end, note that the regression slope estimate $\hat{\beta}_n$ in (2a) converges to $F'(L_p)$ a.s. By comparing (20) and (2a), $c_n - \hat{\beta}_n^{-1} = n(x_n - \bar{x}_n)^2 / \sum^n y_i(x_i-\bar{x}_n)$. Since both procedures converge to L_p for large n, $x_n \approx \bar{x}_n$ and $c_n - \hat{\beta}_n^{-1} \to 0$ follows from the assumption $F'(L_p) > 0$. Therefore the asymptotic optimality of (20) follows from similar results of Anbar (1978) and Lai and Robbins (1981). (Their regularity conditions do not apply directly to the quantal response problem but their technique can be modified to suit our purpose.)

The asymptotic (first order) equivalence of the above two adaptive procedures can be given a more intuitive explanation. In the adaptive RM procedure, the slope $F'(L_p)$ of F at L_p is estimated by the ordinary regression slope estimate $\hat{\beta}_n$. In the adaptive "logit-MLE" procedure, it is estimated via the MLE of the slope parameter λ in the logit model. When x_i are close to L_p, the above proof shows that the two estimates (the latter one being implicit) of $F'(L_p)$ are essentially the same.

6. A simulation study

Under comparison are (i) the logit-MLE version of the sequential design (5) with truncation as defined in (10) (abbreviated as MLE in the Tables), (ii) the adaptive Robbins-Monro (ARM) design with truncation,

$$(21) \qquad x_{n+1} = x_n - \frac{c_n}{n}(y_n - p),\ c_n = \max(\delta, \min(c, \hat{\beta}_n^{-1})),\ c > \delta > 0\ ,$$

where $\hat{\beta}_n$ is defined in (2a), and (iii) the Robbins-Monro (RM) design (2). The Up-and-Down design (4) was also included in the simulation. The results are not reported here since it is consistently the worst.

Note that the use of the logit-MLE design requires the existence and uniqueness of the MLE, condition (9). To facilitate the comparison of the three designs, we start with a common initial design and later branch to the three designs when (9) is satisfied. Two distinct choices of the initial design are considered. The first has fixed design levels and sample size. Initial samples that do not satisfy (9) have to be discarded. The second uses the nonadaptive RM design as the initial design and branches to the logit-MLE and to the ARM at possibly different times. The size of the initial design is random but no simulation sample is discarded. The difference between these two choices of initial design and their practical relevance will be discussed later.

Five models are used in simulating the true quantal response curve, the logit model (6), the probit model (7), the skewed logit model (22),

$$H(x|\lambda) = (1+e^{-\lambda x})^{-2} \quad , \tag{22}$$

the complementary log-log model (23),

$$H(x|\lambda) = 1 - e^{-e^{\lambda x}} \tag{23}$$

and the logit model with a cubic term, $(1 + \exp(-x - \frac{1}{3}x^3)^{-1}$. Results from the last model are not reported in Section 6.1 since they give essentially the same conclusion. For each H, the binary response $y = 0$ or 1 is generated according to $u >$ or $< H(x)$, where u is a uniform random number in $[0,1]$ and x is the corresponding stimulus level. The same set of random numbers u_i is used for all designs under comparison. Note that for the logit-MLE design, the MLE is always computed on the logit assumption, no matter what the true distribution H is.

6.1. Fixed initial designs

A fixed initial design x_i, $i = 1(1)10$, is chosen and the corresponding y_i is generated according to the true distribution H as described above. Let $(\hat{\alpha}_{10},\hat{\lambda}_{10})$ be the MLE of (α,λ) in the logit model based on $\{x_i,y_i\}_1^{10}$. The common starting value for all designs under comparison is chosen to be $x_{11} = \hat{\alpha}_{10} - \hat{\lambda}_{10}^{-1} \ell n(p^{-1} - 1)$ according to (5)(ii). Once x_{11} is chosen, the subsequent design levels $x_{12},\ldots,x_{35}$ are generated according to different design schemes. If the MLE $(\hat{\alpha}_{10},\hat{\lambda}_{10})$ does not exist the simulation sample is discarded. On the other hand, if $(\hat{\alpha}_{10},\hat{\lambda}_{10})$ exists and is unique, the subsequent MLE always exists as is obvious from condition (9). This is repeated for 500 times, including those discarded due to the nonexistence of MLE.

For sample size n, the Monte Carlo mean square error (MSE) of a sequential design is calculated as the average of $(x_n - L_p)^2$ over the simulation samples. In Tables I and II, $\sqrt{MSE}$ are given for the designs (i) - (iii) for estimating $L_{0.5}$ and $L_{0.75}$ for a few initial design. Other initial designs were considered in Wu (1983). The conclusions are very similar. In each table L_p denotes the design level that corresponds to the $100p$ percentile of the true response curve. Therefore, for example, the two initial designs in Tables I(a) and I(b) are identical, but correspond to different percentiles under different response curves.

The results in Tables I and II are summarized as follows.

(A) General comparison of designs

In general, MLE performs substantially better than ARM and RM, with the latter two being quite comparable. Only in Table II(b) does RM-16 (the Robbins-Monro method (2) with $c = 16$) outperform the others. But when the size of the initial design is increased from 10 to 14 as in Table II(b1), MLE has again the best performance.

Within RM we observe the descending order of performance

RM-32 and RM-16 > RM-4 > RM-1 > RM-0.25.

Note that RM-4 is asymptotically optimal for the distributions in Tables I(a)(c), because $F'(L_{0.5}) = 1/4$. RM-4 fails to deliver this asymptotic promise of optimality for n as large as 35. To save space, RM-1 and RM-0.25 are not included in the tables.

TABLE 1. MONTE CARLO $\sqrt{MSE}$ OF SEQUENTIAL DESIGNS FOR ESTIMATING THE 50 PERCENTILE OF THE TRUE QUANTAL RESPONSE CURVE (BASED ON 500 SAMPLES)

I(a) Initial design:

stimulus level	$L_{0.1}$	$L_{0.3}$	$L_{0.5}$	$L_{0.7}$	$L_{0.9}$
no. of observations	1	2	4	2	1

True response curve: logit model (6) with $\alpha = 0$, $\lambda = 1$

design \ n	12	16	20	25	30	35
MLE-30	1.44	1.02	.72	.56	.43	.36
MLE-50	1.40	.78	.53	.47	.40	.36
MLE-100	1.40	.60	.49	.46	.40	.36
MLE-200	1.36	.61	.54	.48	.40	.35
MLE-600	1.48	.78	.56	.45	.41	.36
ARM-16	1.55	1.29	1.09	.92	.78	.67
ARM-30	1.52	1.16	.88	.69	.53	.45
ARM-50	1.54	1.16	.91	.69	.55	.46
ARM-100	1.59	1.24	.99	.80	.62	.51
ARM-600	1.63	2.02	1.66	1.67	1.36	1.12
RM-32	1.84	1.41	1.21	.94	.81	.74
RM-16	1.58	1.31	1.14	.97	.83	.73
RM-4	1.59	1.46	1.37	1.29	1.23	1.19
			M = 114			

where L_p = 100p percentile of the true response curve,

MLE-d = procedure (10) with upper truncation bound d and lower truncation bound 0

ARM-c = procedure (21) with upper truncation bound c and lower truncation bound 0

RM-c = procedure (2) with constant c

M = total number of simulation samples for which no MLE exists

I(b) Initial design (same as I(a)):

stimulus level	$L_{0.3}$	$L_{0.46}$	$L_{0.56}$	$L_{0.66}$	$L_{0.80}$
no. of observations	1	2	4	2	1

True response curve: probit model (7) with $\mu = -0.5$, $\sigma = 3.1915$

design \ n	12	16	20	25	30	35
MLE-30	1.87	1.34	1.07	.85	.82	.77
MLE-50	1.84	1.10	.88	.83	.77	.73
MLE-100	1.95	.93	.85	.74	.83	.62
MLE-200	1.95	.90	.79	.71	.84	.63
MLE-600	1.90	1.13	.87	.75	.76	.63
ARM-16	2.01	1.70	1.54	1.36	1.21	1.08
ARM-30	2.00	1.58	1.42	1.23	1.08	.96
ARM-50	2.06	1.62	1.39	1.21	1.07	.92
ARM-100	2.15	1.80	1.56	1.32	1.19	1.02
ARM-600	2.21	3.32	2.90	2.37	1.86	1.56
RM-32	2.16	1.77	1.51	1.35	1.17	1.05
RM-16	2.01	1.69	1.47	1.28	1.10	.93
RM-4	2.07	1.92	1.81	1.71	1.63	1.56
			M = 56			

I(c) Initial design:

stimulus level	$L_{0.1}$	$L_{0.3}$	$L_{0.5}$	$L_{0.7}$	$L_{0.9}$
no. of observations	1	2	4	2	1

True response curve: skewed logit model (22) with $\lambda = 1$

design	n = 12	16	20	25	30	35
MLE-30	5.17	3.81	2.90	2.05	1.42	1.02
MLE-50	4.93	2.80	1.47	1.00	.93	.89
MLE-100	4.38	1.16	.99	.93	.89	.86
MLE-200	3.61	1.10	.95	.92	.88	.85
MLE-600	4.03	1.57	.90	.89	.85	.82
ARM-16	5.37	4.64	4.12	3.61	3.21	2.88
ARM-30	5.28	4.28	3.61	3.02	2.56	2.17
ARM-50	5.29	4.29	3.63	3.02	2.56	2.17
ARM-100	5.33	4.31	3.65	3.04	2.56	2.18
ARM-600	5.89	4.82	4.11	3.34	2.79	2.34
RM-32	5.22	3.86	2.91	2.05	1.50	1.20
RM-16	5.35	4.64	4.11	3.59	3.20	2.86
RM-4	5.52	5.32	5.17	5.02	4.90	4.80
			M = 112			

I(d) Initial design:

stimulus level	$L_{0.1}$	$L_{0.3}$	$L_{0.5}$	$L_{0.7}$	$L_{0.9}$
no. of observations	1	2	4	2	1

True response curve: complementary log-log model (23) with $\lambda = 1$

design	n = 12	16	20	25	30	35
MLE-30	2.60	1.72	1.25	.86	.67	.54
MLE-50	2.42	1.28	.76	.58	.54	.51
MLE-100	2.09	.96	.65	.59	.52	.49
MLE-200	1.95	.89	.65	.65	.52	.48
MLE-600	2.93	1.38	.87	.62	.71	.47
ARM-16	2.80	2.34	2.03	1.74	1.52	1.34
ARM-30	2.77	2.27	1.93	1.61	1.38	1.18
ARM-50	2.76	2.30	1.96	1.64	1.39	1.19
ARM-100	2.77	2.37	2.00	1.64	1.42	1.20
ARM-600	2.84	3.23	2.62	2.11	1.82	1.51
RM-32	2.81	1.92	1.51	1.11	.95	.84
RM-16	2.79	2.24	1.90	1.59	1.37	1.21
RM-4	2.90	2.72	2.59	2.47	2.38	2.30
			M = 69			

TABLE II. MONTE CARLO $\sqrt{MSE}$ OF SEQUENTIAL DESIGNS FOR ESTIMATING THE 75 PERCENTILE OF THE TRUE QUANTAL RESPONSE CURVE (BASED ON 500 SAMPLES)

II(a) Initial design and true response curve: same as in I(a)

	n					
design	12	16	20	25	30	35
MLE-30	1.43	.87	.70	.61	.56	.48
MLE-50	1.36	.80	.64	.56	.53	.47
MLE-100	1.38	.77	.64	.58	.55	.50
MLE-200	1.43	.77	.65	.59	.56	.50
MLE-600	1.49	.76	.65	.59	.56	.51
ARM-16	1.54	1.19	1.02	.87	.78	.70
ARM-30	1.55	1.21	1.05	.89	.79	.72
ARM-50	1.60	1.26	1.09	.93	.83	.75
ARM-100	1.69	1.37	1.19	1.01	.90	.81
ARM-600	1.71	1.41	1.73	1.39	1.16	.98
RM-32	1.69	1.23	1.16	.93	.87	.78
RM-16	1.51	1.13	.93	.75	.68	.61
RM-4	1.57	1.41	1.28	1.17	1.08	1.01
			M = 114			

For explanation of symbols, see the bottom of Table I(a)

II(b) Initial design and true response curve: same as in I(b)

	n					
design	12	16	20	25	30	35
MLE-30	1.97	1.47	1.21	1.16	1.04	.98
MLE-50	1.95	1.51	1.18	1.12	1.08	1.01
MLE-100	2.09	1.38	1.22	1.15	1.11	1.03
MLE-200	2.33	1.43	1.27	1.14	1.11	1.06
MLE-600	2.33	1.41	1.55	1.16	1.12	1.06
ARM-16	2.02	1.80	1.56	1.40	1.26	1.12
ARM-30	1.98	1.74	1.54	1.32	1.20	1.16
ARM-50	1.99	1.80	1.69	1.46	1.28	1.23
ARM-100	2.07	2.08	1.92	1.63	1.45	1.40
ARM-600	2.08	3.85	3.39	2.98	2.41	2.19
RM-32	2.01	1.52	1.45	1.29	1.19	1.07
RM-16	1.95	1.46	1.20	1.03	.91	.82
RM-4	2.08	1.86	1.71	1.56	1.45	1.36
			M = 56			

II(b1) Initial design:

stimulus level*	$I_{0.3}$	$I_{0.46}$	$I_{0.56}$	$I_{0.66}$	$I_{0.80}$
no. of observations	1	3	6	3	1

Initial sample size: 14

True response curve: same as in II(b)

design	n = 16	20	25	30	35
MLE-30	2.23	1.63	1.37	1.13	.99
MLE-50	2.23	1.42	1.20	1.06	.97
MLE-100	2.27	1.31	1.16	1.07	1.01
MLE-200	2.76	1.40	1.18	1.11	1.03
MLE-600	2.91	1.48	1.17	1.25	1.07
ARM-30	2.24	1.91	1.60	1.44	1.31
ARM-50	2.26	1.97	1.71	1.52	1.38
ARM-100	2.31	2.10	1.82	1.62	1.41
ARM-600	2.96	3.36	3.12	2.72	2.38
RM-32	2.15	1.67	1.36	1.20	1.04
RM-16	2.20	1.77	1.46	1.28	1.13
RM-4	2.31	2.15	2.01	1.90	1.81

*same as in II(b)

II(c) Initial design and true response curve: same as in I(c)

design	n = 12	16	20	25	30	35
MLE-30	3.85	3.04	2.63	2.26	2.02	1.83
MLE-50	3.67	2.59	2.07	1.59	1.25	1.01
MLE-100	3.35	1.80	1.02	.97	.91	.91
MLE-200	2.98	1.09	.96	.96	.91	.89
MLE-600	1.87	1.28	.98	1.17	.91	.90
ARM-30	3.95	3.22	2.76	2.35	2.10	1.88
ARM-50	3.88	2.96	2.36	1.81	1.47	1.20
ARM-100	3.75	2.50	1.76	1.41	1.22	1.11
ARM-200	3.52	2.24	1.87	1.47	1.29	1.15
ARM-600	3.52	2.60	2.53	1.92	1.57	1.33
RM-32	3.92	3.09	2.66	2.25	2.01	1.79
RM-16	4.00	3.48	3.18	2.92	2.72	2.57
RM-4	4.13	3.97	3.85	3.73	3.65	3.58

M = 112

II(d) Initial design and true response curve: same as in I(d)

design	n = 12	16	20	25	30	35
MLE-30	4.63	3.64	3.18	2.70	2.37	2.11
MLE-50	4.48	3.16	2.43	1.79	1.41	1.14
MLE-100	4.36	2.08	1.21	.91	.92	.87
MLE-200	4.67	1.41	1.09	.89	.90	.85
MLE-600	4.97	1.92	1.00	.85	.93	.85
ARM-30	4.74	4.03	3.67	3.35	3.11	2.91
ARM-50	4.77	4.08	3.70	3.37	3.13	2.93
ARM-100	4.89	4.17	3.77	3.43	3.19	2.99
ARM-200	5.16	4.47	4.02	3.66	3.39	3.18
ARM-600	6.54	5.61	4.95	4.45	4.10	3.82
RM-32	4.63	3.75	3.25	2.76	2.40	2.14
RM-16	4.74	4.13	3.81	3.53	3.32	3.14
RM-4	4.89	4.69	4.55	4.43	4.33	4.25
			M = 69			

(B) Superiority of the logit-MLE design.

The superiority of the logit-MLE design (10) with upper truncation bound d and lower truncation bound 0, hereafter denoted as MLE-d, is broad-based. In the nine tables, MLE-50, MLE-100, MLE-200 consistently outperform the best ARM. Except in Table II(c), MLE-30 outperforms the best ARM. The efficiency gain of MLE over ARM is more conspicuous for larger n.

What truncation bound d should be chosen? The MLE designs with $50 \leq d \leq 600$ all perform well. Within this range their difference of performance is probably negligible. MLE-30 does not perform as well, because a forceful truncation like $d = 30$ limits the potential of the MLE design in making more flexible and justifiably large moves.

Since a major purpose for finding better designs is to reduce the number of runs required for satisfying an error bound, we shall measure the efficiency gain of the MLE design over the ARM design by such numbers. In each case, we find the smallest $\sqrt{MSE}$ achieved by the best ARM design at

n = 35. We then find m to be the smallest sample size at which an MLE design achieves the same $\sqrt{MSE}$. In Table III, the values of m are obtained by linear interpolation for the nine tables in Tables I and II. The percentage of runs saved by using the best MLE design instead of the best ARM design ranges from 25% to 57%.

Table III. Values of m for Tables I(a)-(d), II(a)-(d)

I(a)	I(b)	I(c)	I(d)	II(a)	II(b)	II(b1)	II(c)	II(d)
26	16	15	15	18	25	20	16	15

C. ARM or RM?

The best RM design is RM-32 or RM-16 and is quite comparable to the best ARM design. In Tables I(c)(d), II(a)(b)(b1)(d) it even beats the best ARM. But the performance of the RM design depends critically on the choice of the constant c in (2), which may not be available in practical situations. On the other hand, the ARM-c (procedure (21) with upper truncation bound c and lower truncation bound 0) performs well and stably over a broader range of the c values, $16 \leq c \leq 100$ for $L_{0.5}$ and $30 \leq c \leq 200$ for $L_{0.75}$. The ARM-600 design, which uses a loose truncation bound, is consistently worse than the best ARM design and the best RM design. Moreover its MSE exhibits an erratic pattern, e.g., it sometimes increases as n increases. Generally the ARM requires more severe truncation than the MLE. This is because the ARM can make an unduly large move as explained in Section 2.

D. For the same truncation bound, the MLE design always requires more truncations than the ARM Design. It suggests that the MLE design makes large

moves more frequently than the ARM design. Since MLE-100, MLE-200 and MLE-600 do very well in the study, such large moves are probably justifiable.

We have also examined the empirical behavior of the same set of designs for initial designs of size 25. The results are very similar. As the size of the initial design increases, the number of simulation samples for which no MLE exists quickly drops.

6.2 Nonadaptive RM as the initial designs

We choose two starting values $x_1 = L_{0.6}$ and $L_{0.9}$ and three recursive schemes RM-1, RM-4 and RM-16 as the common initial designs. The logit model, (6), the skewed logit model (22) and the complementary log-log model (23), all with $F'(L_{0.5}) = 1/4$, are considered. The two initial designs RM-1 and RM-16 correspond to over- and under-estimates of the true slope $F'(L_{0.5})$. The starting values $L_{0.6}$ and $L_{0.9}$ represent good and poor guesses of $L_{0.5}$.

Define the two switching times as follows:

n_1 = first $n \geq 5$ such that $\hat{\beta}_n$, (2a) is non zero,

n_2 = first $n \geq 5$ such that (9) is satisfied.

For each initial design, denoted by D, three sequential designs are considered:

I. design D for $1 \leq n \leq 35$

II. design D for $1 \leq n \leq n_1$, followed by the adaptive RM, (21), with truncation bounds δ and c for $n_1 + 1 \leq n \leq 35$ (denoted by ARM(δ,c))

III. design D for $1 \leq n \leq n_2$, followed by the logit-MLE, (10), with truncation bounds δ and c for $n_2 + 1 \leq n \leq 35$ (denoted by MLE(δ,c)).

The size of the initial design n_1 or n_2 is <u>random</u> and n_1 is always smaller than or equal to n_2. The time n_1 for switching to the ARM is 5 in most situations. The interquartile range for n_2 (time for switching to

the MLE) is [6,8] or [7,9]. We have also tried a delayed version of design II, namely, to switch to the ARM at time n_2 instead of n_1. Its performance is somewhat inferior and is therefore abandoned. As argued in Section 3, it may not be a good idea to start the logit-MLE recursion as soon as (9) is satisfied. To prevent the MLE estimate from being "trapped" at a point far from $L_{0.5}$, we consider a delayed version of the above design III with lag ℓ,

(24) IV. design III with n_2 replaced by $n_2 + \ell$ (denoted by DMLE (δ,c,ℓ)) .

We choose $(\delta,c) = (0.01,600)$ and $(1,100)$ in the simulation study. Only the estimation of $L_{0.5}$ is considered. The Monte Carlo mean square error of each design is computed based on 1000 simulation samples. To save space, the results on the skewed logit model are not given here since they are quite consistent with those reported in Table IV.

The results in Table IV do not exhibit a clear-cut pattern as those in Section 6.1. To facilitate the following discussion, we group the six initial designs into two categories (G for good, P for poor):

(G) $(L_{0.6}$,RM-1), $(L_{0.6}$,RM-4), $(L_{0.9}$,RM-4)

(P) $(L_{0.6}$,RM-16), $(L_{0.9}$,RM-1), $(L_{0.9}$,RM-16)

Since the performance depends on the initial design, we start our comparison on the nonadaptive RM design.

1) For $x_1 = L_{0.6}$, RM-1 > RM-4 > RM-16 ,

for $x_1 = L_{0.9}$, RM-4 > RM-16 > RM-1 ,

where ">" denotes "better than". The choice of the starting value x_1 interacts with the choice of the constant c in RM-c. Since $F'(L_{0.5}) = 1/4$, RM-4 is asymptotically optimal, which confirms the result for $x_1 = L_{0.9}$. But when the starting value $L_{0.6}$ is close to the true parameter

TABLE IV. MONTE CARLO $\sqrt{MSE}$ OF SEQUENTIAL DESIGNS FOR ESTIMATING THE 50 PERCENTILE OF THE TRUE QUANTAL RESPONSE CURVE (BASED ON 1000 SAMPLES)

IV(a) True response curve: logit model (6) with $\alpha = 0$, $\lambda = 1$ ($F'(L_{0.5}) = 1/4$)

starting value	initial design	sequential design	n = 10	12	16	20	25	30	35
$x_1 = L_{0.6}$	RM-1	RM-1	.46	.45	.43	.41	.39	.38	.36
		ARM(0.01,600)	.53	.52	.50	.48	.63	.64	.58
		ARM(1,100)	.52	.53	.53	.49	.46	.44	.42
		MLE(0.01,600)	.49	.52	.47	.47	.47	.45	.51
		MLE(1,100)	.50	.48	.47	.44	.43	.41	.38
		DMLE(1,100,3)	.50	.59	.46	.45	.41	.41	.39
	RM-4	RM-4	.73	.65	.57	.51	.44	.39	.36
		ARM(0.01,600)	.99	.89	.77	.69	.59	.52	.46
		ARM(1,100)	.99	.89	.77	.68	.59	.51	.46
		MLE(0.01,600)	.80	.77	.72	.64	.60	.59	.55
		MLE(1,100)	.79	.78	.70	.60	.55	.52	.46
		DMLE(1,100,3)	.77	.70	.62	.58	.54	.49	.45
	RM-16	RM-16	1.08	.98	.83	.71	.64	.59	.51
		ARM(0.01,600)	1.11	.94	.75	.62	.54	.50	.43
		ARM(1,100)	1.11	.94	.75	.62	.54	.50	.43
		MLE(0.01,600)	.77	.74	.64	.61	.58	.56	.54
		MLE(1,100)	.77	.69	.61	.56	.51	.47	.44
		DMLE(1,100,3)	.98	.79	.64	.56	.51	.46	.43
$x_1 = L_{0.9}$	RM-1	RM-1	1.29	1.23	1.16	1.10	1.05	1.01	.97
		ARM(0.01,600)	4.44	3.93	3.27	2.83	2.44	2.01	1.65
		ARM(1,100)	1.39	1.26	1.13	1.03	.87	.75	.65
		MLE(0.01,600)	1.27	1.45	1.72	1.31	1.01	.88	.83
		MLE(1,100)	1.17	1.25	1.03	.86	.77	.70	.61
		DMLE(1,100,3)	1.19	1.17	1.01	.87	.76	.70	.61
	RM-4	RM-4	.75	.67	.57	.51	.43	.40	.36
		ARM(0.01,600)	1.07	.95	.75	.67	.61	.60	.54
		ARM(1,100)	1.07	.94	.74	.64	.55	.48	.43
		MLE(0.01,600)	.79	.95	.81	.65	.60	.59	.58
		MLE(1,100)	.78	.77	.67	.62	.54	.51	.47
		DMLE(1,100,3)	.75	.87	.64	.58	.51	.47	.45
	RM-16	RM-16	1.07	.97	.80	.72	.63	.58	.51
		ARM(0.01,600)	1.00	.84	.67	.59	.50	.46	.41
		ARM(1,100)	1.00	.84	.67	.59	.50	.46	.41
		MLE(0.01,600)	1.09	1.01	.94	.92	.89	.86	.84
		MLE(1,100)	1.06	.97	.87	.79	.74	.68	.62
		DMLE(1,100,3)	1.09	.92	.74	.65	.57	.52	.48

where L_p = 100p percentile of the true response curve

RM-c = procedure (2) with constant c

ARM(δ,c) = procedure (21) with lower and upper truncation bounds δ and c

MLE(δ,c) = procedure (10) with lower and upper truncation bounds δ and c

DMLE(δ,C,ℓ) = delayed MLE procedure (24) with lower and upper trunction bounds δ and c, lag of delay ℓ

IV(b) True response curve: complementary log-log model (23) with $\lambda = 0.721$ ($F'(L_{0.5}) = 1/4$)

starting value	initial design	sequential design	n = 10	12	16	20	25	30	35
$x_1 = L_{0.6}$	RM-1	RM-1	.44	.43	.40	.39	.37	.35	.34
		ARM(0.01,600)	.51	.50	.48	.46	.62	.63	.57
		ARM(1,100)	.50	.51	.51	.48	.45	.43	.40
		MLE(0.01,600)	.47	.50	.45	.45	.46	.42	.50
		MLE(1,100)	.47	.45	.45	.42	.41	.38	.36
		DMLE(1,100,3)	.48	.57	.44	.43	.39	.39	.36
	RM-4	RM-4	.69	.63	.55	.50	.43	.39	.35
		ARM(0.01,600)	.97	.88	.77	.68	.58	.52	.47
		ARM(1,100)	.97	.88	.77	.68	.58	.52	.47
		MLE(0.01,600)	.73	.70	.65	.82	.57	.55	.62
		MLE(1,100)	.71	.70	.63	.57	.53	.49	.45
		DMLE(1,100,3)	.74	.68	.62	.57	.52	.47	.44
	RM-16	RM-16	1.09	.97	.82	.72	.64	.58	.51
		ARM(0.01,600)	1.09	.94	.75	.61	.54	.50	.43
		ARM(1,100)	1.09	.94	.75	.61	.54	.50	.43
		MLE(0.01,600)	.76	.71	.62	.60	.57	.55	.53
		MLE(1,100)	.75	.69	.59	.55	.50	.47	.44
		DMLE(1,100,3)	.98	.75	.60	.54	.48	.45	.42
$x_1 = L_{0.9}$	RM-1	RM-1	.86	.82	.76	.72	.68	.65	.63
		ARM(0.01,600)	3.84	3.41	2.83	2.45	2.01	1.75	1.41
		ARM(1,100)	1.11	1.03	.97	.88	.76	.68	.61
		MLE(0.01,600)	.87	1.03	1.36	.91	.83	.66	.61
		MLE(1,100)	.81	.85	.73	.67	.58	.55	.51
		DMLE(1,100,3)	.83	.89	.77	.66	.60	.54	.49
	RM-4	RM-4	.66	.60	.53	.47	.41	.37	.34
		ARM(0.01,600)	.95	.85	.75	.64	.55	.48	.43
		ARM(1,100)	.95	.85	.71	.62	.53	.47	.42
		MLE(0.01,600)	.71	.78	.74	.61	.54	.51	.48
		MLE(1,100)	.70	.69	.63	.58	.50	.47	.43
		DMLE(1,100,3)	.73	.83	.60	.54	.50	.45	.42
	RM-16	RM-16	1.04	.96	.84	.73	.62	.57	.51
		ARM(0.01,600)	1.01	.89	.70	.61	.51	.46	.41
		ARM(1,100)	1.01	.89	.70	.61	.51	.46	.41
		MLE(0.01,600)	.94	.83	.75	.72	.69	.67	.64
		MLE(1,100)	.94	.83	.71	.64	.59	.53	.50
		DMLE(1,100,3)	1.06	.86	.67	.59	.52	.48	.45

$L_{0.5}$, the simulation result defies the asymptotic prediction. In fact, according to a standard asymptotic result on RM (Lai and Robbins, 1979, Theorem 2 (ii)), the convergence rate of $x_n - L_{0.5}$ for x_n generated by RM-1 is of order $n^{-0.25}$ while both RM-4 and RM-16 give the better convergence rate $n^{-0.5}$. The reason that the asymptotic results are not applicable here is because, for x_1 close to $L_{0.5}$, a small c in the RM recursion (2) is needed to ensure a steady convergence to $L_{0.5}$. Even a moderate value like $c = 4$ will make the correction

$$|x_{n+1} - x_n| = \frac{4}{n}\frac{1}{2} = \frac{2}{n}$$

too fluctuating for small n.

2) The ARM designs, despite the asymptotic promise, do not do as well as the RM-designs. For the poor initial design $(L_{0.9}$,RM-1), ARM(0.01,600) performs miserably. By choosing tighter truncation bounds, ARM(1,100) improves over ARM(0.01,600) and even beats RM-1 in the standard logit model. The only other case that gives the ARM an edge over the RM is the initial design $(L_{0.9}$,RM-16) in category (P).

3) Three versions of the MLE designs are under comparison. There is a substantial improvement over MLE(0.01,600) by using MLE(1,100) with tighter truncation bounds. Additional improvement is made by using the delayed-MLE design DMLE(1,100,3) with lag 3. When (and only when) the initial designs are in category (P), MLE(1,100) and DMLE(1,100,3) beat the RM design. In a few cases, MLE(0.01,600) also beats the RM-design. The superior performance of the RM designs in category (G) depends critically on good prior knowledge of $L_{0.5}$ and $F'(L_{0.5})$. When such knowledge is not available, the MLE design, MLE(1,100), does better. Of course the choice of the tighter bounds 1 and 100 in MLE(1,100) assumes a good knowledge of the slope $F'(L_{0.5})$ (but not of $L_{0.5}$), though not in the same degree as the RM designs.

4) The conclusions seem to be independent of the choice of distributions. The results from (not reported here) the skewed logit model lead to the same conclusions.

7. Concluding remarks

In this paper a new class of sequential designs for binary response data is proposed. Its consistency and asymptotic normality, via its connection with the Robbins-Monro method, are demonstrated under rather restrictive conditions. These methods are compared in a simulation study for sample sizes up to 35. It is somewhat unexpected that their relative performance depends quite heavily on the choice of the initial designs. The fixed initial designs in Section 6.1 have design levels spreading evenly over wider intervals. The levels of the nonadaptive Robbins-Monro designs in Section 6.2 tend to be unevenly distributed and not so wide-spread.

The empirical results suggest that, when a good initial RM design is available, the RM design should be used for the first phase of the experiment. For larger n, RM may be replaced by ARM or MLE to take advantage of the asymptotic optimality of the latter designs. But if the quality of the initial RM design is not certain, a MLE design with tighter truncation bounds or its delayed version should be used. It is possible that other modifications of the MLE design will further enhance its utility. This merits further study.

It is easy to conceive practical situations in which other initial designs are preferred. The experimenter may not have any vague idea about L_p and $F'(L_p)$, two elements critical to the performance of the RM design. One common practice is to choose a wide interval that is believed to contain the target value L_p, and to place the initial design levels evenly over the

interval, including the two end points. It avoids the adverse effect of extremely misleading guesses. The business of choosing the first five to eight design levels is a very subjective one. A good experimenter will exercise his best judgement and utilize whatever prior knowledge available to him in making his choice.

The simulation results of Section 6.1 suggest that the MLE design can take full advantage of the past information if the initial design levels are wide-spread, and the response region and the nonresponse region overlap. The latter condition implies that (9) is satisfied. Since the initial samples that do not satisfy (9) are discarded in the simulation, our conclusion should only apply to those initial samples that are ready for the application of the MLE design. For initial sample size 10, the number of discarded samples is not negligible (between 56 and 114 out of 500, Table II). When the initial sample size is raised to 14, as in Table II(b1), the number drops from 56 to 14 while the number of runs saved by the latter design increases (Table III). The percentage of discarded samples would be much smaller in practical situations since any sensible experimenter should be able to conduct the first ten runs to satisfy condition (9).

In summary, the proposed MLE designs are useful alternatives to the standard ones. They perform well in some selected situations. Further study is needed to find ways of improving their efficiency and to identify situations, including the choice of initial designs, in which they excel over their competitors.

References

Abdelhamid, S. N. (1973), "Transformations of observations in stochastic approximation," Annals of Statistics, 1, 1158-1174.

Anbar, D. (1973), "On optimal estimation methods using stochastic approximation procedures," Annals of Statistics, 1, 1175-1184.

_____ (1978), "A stochastic Newton-Raphson method," Journal of Statistical Planning and Inference, 2, 153-163.

Berkson, J. (1955), "Maximum likelihood and minimum χ^2 estimates of the logistic function," Journal of the American Statistical Association, 50, 130-162.

Chambers, E. A. and Cox, D. R. (1967), "Discrimination between alternative binary response models," Biometrika, 54, 573-578.

Cox, D. R. (1970), Analysis of Binary Data, London: Methuen.

Chung, K. L. (1954), "On a stochastic approximation method," Annals of Mathematical Statistics, 25, 463-483.

Dixon, W. J. and Mood, A. M. (1948), "A method for obtaining and analyzing sensitivity data," Journal of the American Statistical Association, 43, 109-126.

Dubins, L. E. and Freedman, D. A. (1965), "A sharper form of the Borel-Cantelli Lemma and the strong law," Annals of Mathematical Statistics, 36, 800-807.

Finney, D. J. (1978), Statistical Method in Biological Assay, London: Griffin.

Freeman, P. R. (1970), "Optimal Bayesian sequential estimation of the median effective dose," Biometrika, 57, 79-89.

Hodges, J. L. and Lehmann, E. L. (1955), "Two approximations to the Robbins-Monro process," Proceedings of the 3rd Berkeley Symposium, 1, 95-104.

Lai, T. L. and Robbins, H. (1979), "Adaptive design and stochastic approximation," Annals of Statistics, 7, 1196-1221.

_____ (1981), "Consistency and asymptotic efficiency of slope estimates in stochastic approximation schemes," Z. Wahrscheinlichkeitstheorie verw. Gebiete, 56, 329-360.

Leonard, T. (1982), "An inferential approach to the bioassay design problem," Technical Summary Report #2416, Mathematics Research Center, University of Wisconsin-Madison.

Lord, F. M. (1971), "Tailored testing, an application of stochastic approximation," Journal of the American Statistical Association, 66, 707-711.

Miller, R. G. and Halpern, J. W. (1980), "Robust estimators for quantal bioassay," Biometrika, 67, 103-110.

Owen, R. J. (1975), "A Bayesian sequential procedure for quantal response in the context of adaptive mental testing," Journal of the American Statistical Association, 70, 351-356.

Robbins, H. and Monro, S. (1951), "A stochastic approximation method," Annals of Mathematical Statistics, 29, 400-407.

Robbins, H. and Siegmund, D. (1971), "A convergence theorem for non-negative almost sure supermartingale and some applications," in Optimization Methods in Statistics (ed. J. N. Rustagi), Academic Press, 233-257.

Rose, R. M., Teller, D. Y. and Rendleman, P. (1970), "Statistical properties of staircase estimates," Perception and Psychophysics, 8, 199-204.

Sacks, J. (1958), "Asymptotic distribution of stochastic approximation procedures," Annals of Mathematical Statistics, 29, 373-405.

Silvapulle, M. J. (1981), "On the existence of maximum likelihood estimators for the binomial response model," Journal of the Royal Statistical Society B, 43, 310-313.

Tsutakawa, R. K. (1972), "Design of experiment for bioassay," Journal of the American Statistical Association, 67, 584-590.

Wetherill, G. B. (1963), "Sequential estimation of quantal response curves," (with Discussions) Journal of the Royal Statistical Society B, 25, 1-48.

_____ (1966), Sequential Methods in Statistics, London: Methuen.

Wu, C. F. J. (1983), "Efficient model-based sequential designs for sensitivity experiments," Technical Summary Report #2563, Mathematics Research Center, University of Wisconsin-Madison.

CFJW/jvs/ed

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: #2692	2. GOVT ACCESSION NO.: AD-A142 904
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Efficient Sequential Designs with Binary Data	5. TYPE OF REPORT & PERIOD COVERED: Summary Report - no specific reporting period
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): C. F. Jeff Wu	8. CONTRACT OR GRANT NUMBER(s): DAAG29-82-K0154, DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Work Unit Number 4 - Statistics & 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: May 1984
	13. NUMBER OF PAGES: 37
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)

Logit, Optimal design, Quantal response curve, Robbins-Monro stochastic approximation, Sensitivity experiments, Spearman-Kärber estimator, Up-and-Down method

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

A class of sequential designs for estimating the percentiles of a quantal response curve is proposed. Its updating rule is based on an efficient summary of all the data available via a parametric model. The "logit-MLE" version of the proposed designs can be viewed as a natural analogue of the Robbins-Monro procedure in the case of binary data. It is shown to be asymptotically consistent, distribution-free and optimal via its connection with the latter procedure. For certain choices of initial designs the proposed method performs very well in a simulation study for sample sizes up to 35. A nonparametric sequential design, via the Spearman-Kärber estimator, for estimating the median is also proposed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)