Optimal Bayesian Sequential Estimation of the Median Effective Dose

P. R. Freeman

Biometrika, Volume 57, Issue 1 (Apr., 1970), 79-89.

Your use of the JSTOR database indicates your acceptance of JSTOR's Terms and Conditions of Use. A copy of JSTOR's Terms and Conditions of Use is available at http://www.jstor.org/about/terms.html, by contacting JSTOR at jstor-info@umich.edu, or by calling JSTOR at (888)388-3574, (734)998-9101 or (FAX) (734)998-9113. No part of a JSTOR transmission may be copied, downloaded, stored, further transmitted, transferred, distributed, altered, or otherwise used, in any form or by any means, except: (1) one stored electronic and one paper copy of any article solely for your personal, non-commercial use, or (2) with prior written permission of JSTOR and the publisher of the article or other text.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Biometrika is published by Biometrika Trust. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/bio.html.

Biometrika
©1970 Biometrika Trust

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2000 JSTOR

http://www.jstor.org/
Wed Mar 15 22:11:46 2000

Biometrika (1970), **57**, 1, *p.* 79
With 4 text-figures
Printed in Great Britain

Optimal Bayesian sequential estimation of the median effective dose

BY P. R. FREEMAN
University College London

SUMMARY

This paper concerns the sequential design of experiments for estimating the median lethal dose parameter of a quantal, logistic dose response curve. Bayesian decision theory is used to provide a stopping rule and terminal decision rule to minimize the prior expectation of the total cost of observation plus estimation loss. Special cases in which trials can only be performed at one, two or three dose levels are considered, and optimal strategies are evaluated numerically. The expected losses are compared with those of the up and down method using a fixed sample size equal to the prior expectation of the number of trials under the sequential design. The up and down method is shown to be surprisingly good, having efficiency usually in excess of 90 %.

1. INTRODUCTION

Many techniques exist for estimating the parameters of the logistic model

$$p = \frac{1}{1+e^{-\beta(x-\gamma)}}, \tag{1·1}$$

where the independent variable x is here called the dose level, and the response is a quantal random variable $\tilde{y}$ such that

$$\text{prob}\,(\tilde{y}=1) = p, \quad \text{prob}\,(\tilde{y}=0) = 1-p. \tag{1·2}$$

Both the 'up and down' method of Dixon & Mood (1948) and the 'stochastic approximation' method of Robbins & Monro (1951) have been extensively explored; see e.g. Wolfowitz (1952), Brownlee, Hodges & Rosenblatt (1953), Chung (1954), Sacks (1958) and Kesten (1958). It is known, for example, that for estimating γ the up and down method is 141 % efficient relative to the usual logistic method using five equally spaced groups centred exactly on $x = \gamma$ with optimal spacing, and is 80 % efficient relative to the best procedure of all, which concentrates all observations at $x = \gamma$. Although both these methods have the great advantage of simplicity, the sequence of dose levels they prescribe, especially for small sample sizes, may be suboptimal, and several modifications (Cochran & Davis, 1963–5; Wetherill, 1963; Wetherill, Chen & Vasudeva, 1966) have been suggested. Moreover, the choice of stopping rule is left to the experimenter, so that the payoff between the cost of further observation and that of less accurate estimation is not considered explicitly.

Marks (1962) first considered the problem as one of Bayesian sequential design, but obtained detailed results only in the special cases when the prior for γ is a two-point distribution

or when the dose response curve is a step function. Because of the analytical intractability of the likelihood function

$$\frac{\exp\left\{-\beta \sum_{i=1}^{n} (1-y_i)(x_i-\gamma)\right\}}{\prod_{i=1}^{n} \{1+e^{-\beta(x_i-\gamma)}\}} \tag{1·3}$$

when observations $y_1, \ldots, y_n$ are made at dose levels $x_1, \ldots, x_n$, attention has had to be confined in the present paper to other special cases, when there are only one, two or three dose levels available for experimentation. Another defect which this approach shares with previous ones is that the slope parameter β is assumed known, at least in combination with certain other quantities defined later. It is clearly possible to overcome this by starting with a joint prior for both parameters and using this throughout, but this introduces extra complexity and has not been attempted here.

2. Experimenting at one dose level

Suppose that only one dose level x is available for use. For reasons to appear shortly, x is to be thought of as given by outside conditions and not at the choice of the experimenter. This highly simplified situation obviously resolves itself into one of simple binomial sampling, with p given by (1·1) being the probability of response, conditional on γ. If the prior distribution for γ is taken from the conjugate family

$$f(\gamma|r_0, n_0) \propto \frac{e^{r_0\beta(x-\gamma)}}{\{1+e^{\beta(x-\gamma)}\}^{n_0}}, \tag{2·1}$$

this corresponds to the usual beta prior $\propto p^{r_0-1}(1-p)^{n_0-r_0-1}$ for p.

If r responses are obtained in n independent trials, the posterior for γ has the same form as the prior with parameters (r_0, n_0) replaced by (r_0+r, n_0+n). The process of sequential experimentation can thus be represented as a random walk in the (r, n) plane starting at the point (r_0, n_0). Since the following analysis requires that relevant distributions be proper, we take without loss of generality $r_0 = 1$, $n_0 = 2$.

Let the cost of each observation be c, assumed independent of the actual outcome, and $k(\hat{\gamma}-\gamma)^2$ be the loss incurred by taking $\hat{\gamma}$ as an estimate of γ. This corresponds to the loss function

$$\frac{k}{\beta^2}\left[\log\left\{\frac{p(1-\hat{p})}{\hat{p}(1-p)}\right\}\right]^2$$

for p. At a general point (r, n) let (Lindley & Barnett, 1965)

$D(r, n)$ denote the loss incurred by stopping and using the optimal estimate for γ,

$B^*(r, n)$ denote the loss incurred by taking one further observation and using the optimal strategy thereafter,

$B(r, n)$ denote the loss incurred by using the optimal strategy.

The dynamic programming equations are

$$B^*(r, n) = c + \frac{r}{n}B(r+1, n+1) + \frac{n-r}{n}B(r, n+1), \tag{2·2}$$

$$B(r, n) = \min\{B^*(r, n), D(r, n)\}, \tag{2·3}$$

$$D(r, n) = \frac{k}{\beta^2}\left\{\sum_{i=r}^{\infty}\frac{1}{i^2} + \sum_{i=n-r}^{\infty}\frac{1}{i^2}\right\} \quad (1 \leqslant r \leqslant n-1, \quad n \geqslant 2). \tag{2·4}$$

The last of these quantities is k times the posterior variance of γ at the point (r, n). These equations thus determine the optimal strategy which takes the form of a division of the integer lattice of the (r, n) plane into continuation and stopping regions.

By methods similar to those of B. N. Barnett's unpublished thesis, it is possible to show that for a given n there exist numbers $r_1(n)$ and $r_2(n)$ such that points (r, n) are continuation points for $r = 1, \ldots, r_1(n)$ and $r = r_2(n), \ldots, n-1$ and stopping points for $r = r_1(n)+1, \ldots, r_2(n)-1$. The symmetry equation $D(n-r, n) = D(r, n)$ in fact ensures that $r_1(n)+r_2(n) = n$. As n increases $r_1(n)$ decreases and ultimately becomes 1. The smallest value of n for which all points (r, n) are stopping points for $r = 1, \ldots, n-1$ is $[k/(c\beta^2)]+2$. It is thus a simple numerical matter to work backwards from this value of n using equations (2·2) to (2·4) recursively.

Graphs of the boundaries between the continuation and stopping regions are given in Fig. 1 for $c\beta^2/k = 0{\cdot}05$, $0{\cdot}005$, $0{\cdot}0005$. Other cases for $c\beta^2/k$ between $0{\cdot}0001$ and $0{\cdot}01$ have been computed. The symmetry is displayed more clearly by plotting $r-\frac{1}{2}n$, rather than r, against n. The shapes of the curves reflect the fact that when x is close to γ, so that r remains close to $\frac{1}{2}n$, the experiment is much more informative than when x is farther away. This is in sharp contrast with the sequential estimation of p using quadratic loss $k(\hat{p}-p)^2$ (Ferguson, 1967, pp. 318–324). Here it is obviously least informative to observe an evenly balanced sequence of responses and non-responses, so that the largest sample sizes are needed when $r = \frac{1}{2}n$ and the smallest when $r = 0$ or $r = n$. Optimal boundaries are shown in Fig. 2 for various values of c/k.

It should perhaps again be emphasized that these boundaries are the optimal ones whatever prior distribution is taken for γ, provided it belongs to the conjugate family (2·1). The higher the values of r_0 and n_0 taken, the greater is the prior knowledge about γ, the further 'up and to the right' is the starting point of the random walk in the (r, n) plane, and hence the smaller is the expected sample size. An improper prior, however, corresponds to a starting point 'below and to the left of' the point (1, 2), so that there is positive probability of missing the stopping boundary altogether.

3. Experimenting at two dose levels

Once trials are allowed to take place at either of two dose levels, the problem literally takes on a new dimension. At each step, there are three alternative actions: 'stop and estimate γ', 'perform one trial at dose level x_1' and 'perform one trial at dose level x_2'.

The likelihood function when r_1 responses are obtained from n_1 independent trials at x_1 and r_2 responses from n_2 trials at x_2 is

$$f(\gamma|r_1, r_2, n_1, n_2) \propto \frac{e^{\beta r_1(x_1-\gamma)}}{\{1+e^{\beta(x_1-\gamma)}\}^{n_1}} \frac{e^{\beta r_2(x_2-\gamma)}}{\{1+e^{\beta(x_2-\gamma)}\}^{n_2}}. \tag{3·1}$$

Let

$$\theta = e^{\beta(x_2-x_1)}, \quad \text{assumed less than one so that } x_1 > x_2 \text{ if } \beta > 0, \tag{3·2}$$

$$I_m(r, n_1, n_2) = \int_{-\infty}^{\infty} \frac{e^{ru}(\log u)^m}{(1+e^u)^{n_1}(1+\theta e^u)^{n_2}} du \quad (m = 0, 1, 2), \tag{3·3}$$

so that

$$f(\gamma|r, n_1, n_2) = \frac{\beta}{I_0(r, n_1, n_2)} \frac{e^{r\beta(x_2-\gamma)}}{\{1+e^{\beta(x_1-\gamma)}\}^{n_1}\{1+\theta e^{\beta(x_1-\gamma)}\}^{n_2}}, \tag{3·4}$$

where $r = r_1+r_2$ is the sufficient statistic for γ.

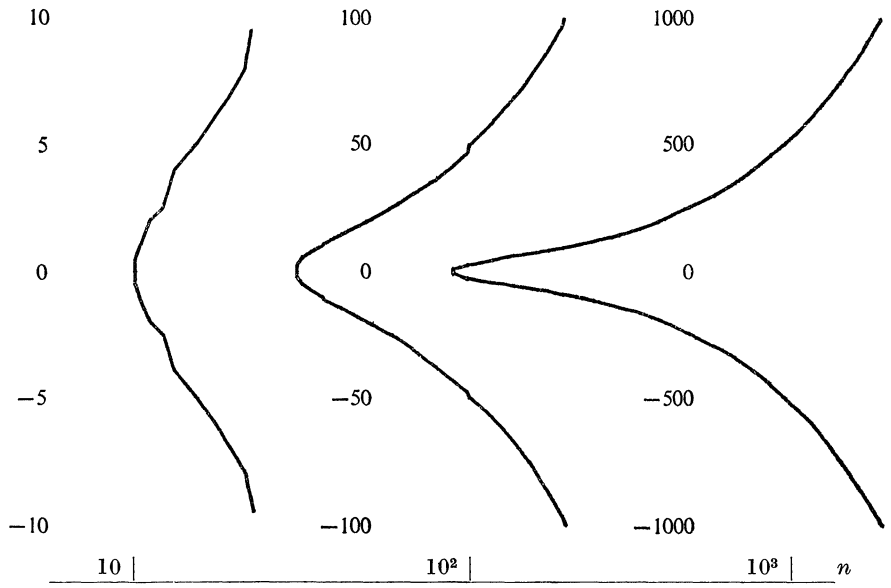

Fig. 1. Optimal stopping boundaries for experimentation at one dose level, with $r-\frac{1}{2}n$ plotted against $\log n$; $c\beta^2/k = 0{\cdot}05,\ 0{\cdot}005,\ 0{\cdot}0005$.

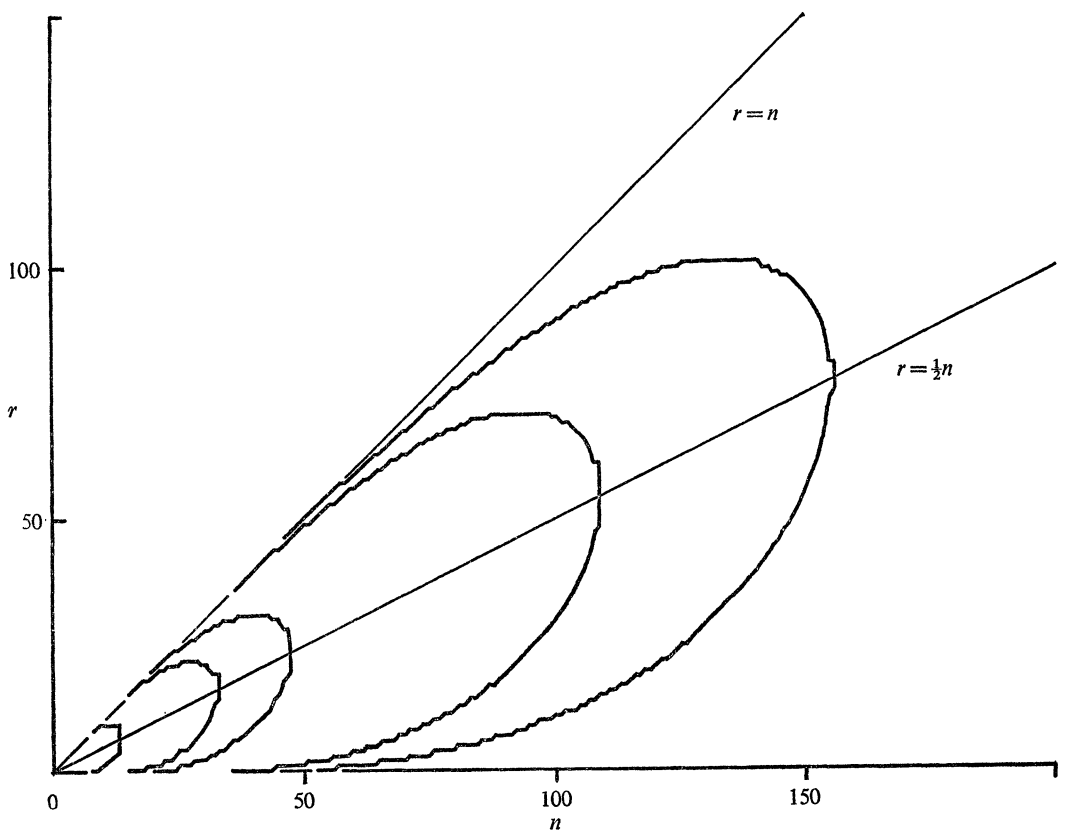

Fig. 2. Optimal stopping boundaries for estimation of a binomial parameter p with quadratic loss and with r plotted against n; $c/k = 0{\cdot}001,\ 0{\cdot}0002,\ 0{\cdot}0001,\ 0{\cdot}00002,\ 0{\cdot}00001$.

The process can thus be represented in (r, n_1, n_2) space. Possible transitions are to $(r+1, n_1+1, n_2)$ or (r, n_1+1, n_2) if a trial is performed at x_1, or to $(r+1, n_1, n_2+1)$ or (r, n_1, n_2+1) if a trial at x_2.

The equations analagous to (2·2) to (2·4) are

$$B_1^*(r, n_1, n_2) = c + \frac{I_0(r+1, n_1+1, n_2)}{I_0(r, n_1, n_2)} B(r+1, n_1+1, n_2) + \left\{1 - \frac{I_0(r+1, n_1+1, n_2)}{I_0(r, n_1, n_2)}\right\} B(r, n_1+1, n_2), \quad (3{\cdot}5)$$

$$B_2^*(r, n_1, n_2) = c + \theta \frac{I_0(r+1, n_1, n_2+1)}{I_0(r, n_1, n_2)} B(r+1, n_1, n_2+1) + \left\{1 - \theta \frac{I_0(r+1, n_1, n_2+1)}{I_0(r, n_1, n_2)}\right\} B(r, n_1, n_2+1), \quad (3{\cdot}6)$$

$$B(r, n_1, n_2) = \min\{D(r, n_1, n_2), \quad B_1^*(r, n_1, n_2), \quad B_2^*(r, n_1, n_2)\}, \quad (3{\cdot}7)$$

$$D(r, n_1, n_2) = \frac{k}{\beta^2}\left[\frac{I_2(r, n_1, n_2)}{I_0(r, n_1, n_2)} - \left\{\frac{I_1(r, n_1, n_2)}{I_0(r, n_1, n_2)}\right\}^2\right]. \quad (3{\cdot}8)$$

Since it has proved impossible to find closed analytic expressions for the I integrals, all subsequent work has used their numerical evaluation by means of Gauss–Laguerre quadrature. In the first stage, all points $N = n_1 + n_2 = 2(1)\,25$, $r = 1(1)\,N-1$ for $\theta = 0{\cdot}4$ and $\theta = 0{\cdot}1$ were systematically worked through. These values of θ correspond to probabilities of response $(0{\cdot}38, 0{\cdot}62)$ and $(0{\cdot}24, 0{\cdot}76)$ respectively at the two dose levels when they are placed symmetrically around the true value of γ. It was felt that these spacings represented the sort of values that might well be encountered in practice. A value of 0·01 was later tried for θ, corresponding to probabilities $(0{\cdot}091, 0{\cdot}909)$ but this caused problems in the accuracy of the quadrature for moderate values of N, so that only small sample sizes could be used.

The symmetry relations

$$I_0(n_1+n_2-r, n_2, n_1) = \theta^{r-n_1} I_0(r, n_1, n_2), \quad D(n_1+n_2-r, n_2, n_1) = D(r, n_1, n_2)$$

were useful in reducing the amount of computing.

For various values of $c\beta^2/k$ sufficiently large to ensure that sampling could not continue beyond $N = 24$, equations (3·5) to (3·8) were applied recursively to partition the space into regions corresponding to the three possible actions. The optimal sequential strategy could then be determined from these partitions for any particular prior for γ, that is any starting point in the space. Such a strategy is shown in Fig. 3, where the digit at each branch point indicates the dose level at which the next trial is to be performed. If a response is then observed, the path to the right is taken in order to determine the next dose level, while a non-response causes the path to the left to be followed. When two possible paths through the tree give rise to the same posterior distribution for γ, they are merged since it is only this 'current' posterior that can influence the choice of action for the next trial. It is evident from Fig. 3 how a less informative sequence of observations necessitates a larger sample size. For example, we need to make ten trials at x_1, the higher dose level, if these all turn out to yield non-responses, while we may stop after five trials at x_1 if we obtain two or three responses. Similarly three responses out of four trials at x_1 and two non-responses out of two trials at x_2 enable us to stop after six trials.

On having discovered the general shapes of the three regions in (r, n_1, n_2) space, it was then possible to write a computer program to explore the boundaries between them without the necessity of investigating the entire space. This allowed much larger sample sizes and hence smaller values of $c\beta^2/k$ to be used. As Lindley & Barnett (1965) found, the continuation regions contained very elongated sections so that the actual sample sizes are only

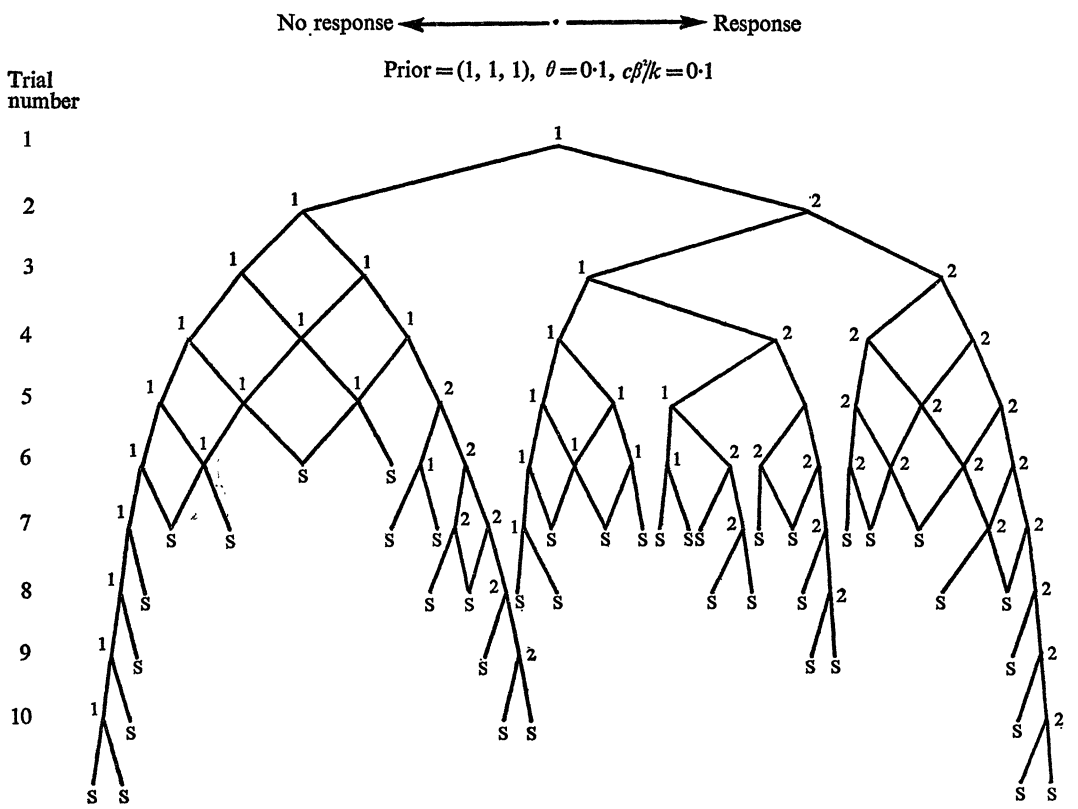

Fig. 3. Optimal sequential schemes using two dose levels. The digit at each branch point indicates the dose level to be used for the next trial. If a response is observed the branch to the right is taken, if a non-response then the branch to the left.

small fractions of their maximum possible values. In Fig. 3, for example, the x_1 continuation region stretches out as far as the point (1, 0, 76), and similarly the x_2 region to (75, 76, 0), while the optimal policy illustrated has a maximum sample size of ten.

As $c\beta^2/k$ decreases, the optimal policies rapidly become tedious to write out in full, but it remains a simple matter after each trial to check in which region the current position of the random walk lies.

4. EXPERIMENTING AT THREE DOSE LEVELS

For three dose levels $x_3 < x_2 < x_1$, assumed equally spaced, the theory now gives a density for γ

$$f(\gamma|r, n_1, n_2, n_3) = \frac{\beta}{I_0(r, n_1, n_2, n_3)} \frac{e^{r\beta(x_1-\gamma)}}{\{1+e^{\beta(x_1-\gamma)}\}^{n_1}\{1+\theta\, e^{\beta(x_1-\gamma)}\}^{n_2}\{1+\theta^2\, e^{\beta(x_1-\gamma)}\}^{n_3}}, \quad (4{\cdot}1)$$

where
$$I_m(r, n_1, n_2, n_3) = \int_{-\infty}^{\infty} \frac{e^{ru}(\log u)^m \, du}{(1+e^u)^{n_1}(1+\theta\, e^u)^{n_2}(1+\theta^2 e^u)^{n_3}} \quad (m = 0, 1, 2), \tag{4·2}$$

$$\theta = e^{\beta(x_2-x_1)} = e^{\beta(x_3-x_2)} \tag{4·3}$$

and $r = r_1+r_2+r_3$ is again sufficient for γ.

The dynamic programming equations are obvious generalizations of (3·5) to (3·8). Again, numerical investigation is necessary and with four dimensions now the amount of computation rapidly becomes prohibitive. The continuation regions for

$$N = n_1+n_2+n_3 = 2(1)\,12, \quad r = 1(1)\,N-1, \quad \theta = 0{\cdot}4$$

have been calculated directly, but to go even to $N = 15$ would require nearly four times as much computing time. An example of the resulting policies is shown in Fig. 4. This is, in fact, likely to be slightly suboptimal, since after ten trials one runs up against the upper limit of $N = 12$, so that the two outermost branches of the tree have been truncated. The same considerations as before apply to the information content of particular observations and sample size needed before stopping.

5. Comparison with the up and down method

Given that we are unable, as yet, to contemplate evaluating the optimal policy for, say, up to fifty trials over twenty dose levels, we can perhaps say something useful about the general efficiency of commonly used techniques, and of the up and down method in particular, in the relatively simple cases where we know what the most efficient policy is. The obvious difficulty in getting any sort of comparison between a Bayesian policy and a non-Bayesian one is that the former is necessarily formulated in more detail. The up and down method provides no stopping rule and, although the experimenter has to use his initial ideas about γ in order to choose the dose level for his first trial, he does not have to explicitly specify his prior distribution. To make any sort of comparison possible, therefore, we must add the necessary extra specification of exactly how the up and down method is to be used. Although there are many possible ways of doing this, the following is perhaps the most straightforward.

Suppose experimenter A has a certain prior for γ and a certain value of $c\beta^2/k$, and that he conducts trials sequentially according to the optimal Bayesian scheme as found above. By suitably weighting every possible path through the decision tree with its corresponding probability, he can obtain his prior expectation of the sample size and of the overall loss, including both the cost of the observations and the terminal estimation loss. Now suppose that experimenter B has the same prior for γ and the same value of $c\beta^2/k$, and that he conducts trials sequentially according to the up and down method, using the same set of dose levels as A and the same initial dose level for his first trial. If we further suppose that he uses a fixed sample size equal to the integer nearest to A's prior expected sample size, we may then evaluate his expected loss and define the relative efficiency of the two schemes as the ratio of their expected losses.

While the up and down method does not, of course, necessarily require a fixed sample size, this seems the least arbitrary stopping rule that can be used here for purposes of comparison. It might be fairer to compare this with the optimal Bayesian scheme which is also constrained to be of fixed sample size, but the expected loss of this latter policy must

obviously lie between that of the fixed sample size up and down method and that of the true optimal Bayesian policy. The efficiency of the up and down method with respect to it would thus be greater than that with respect to the true optimum which are calculated here, although still less than unity.

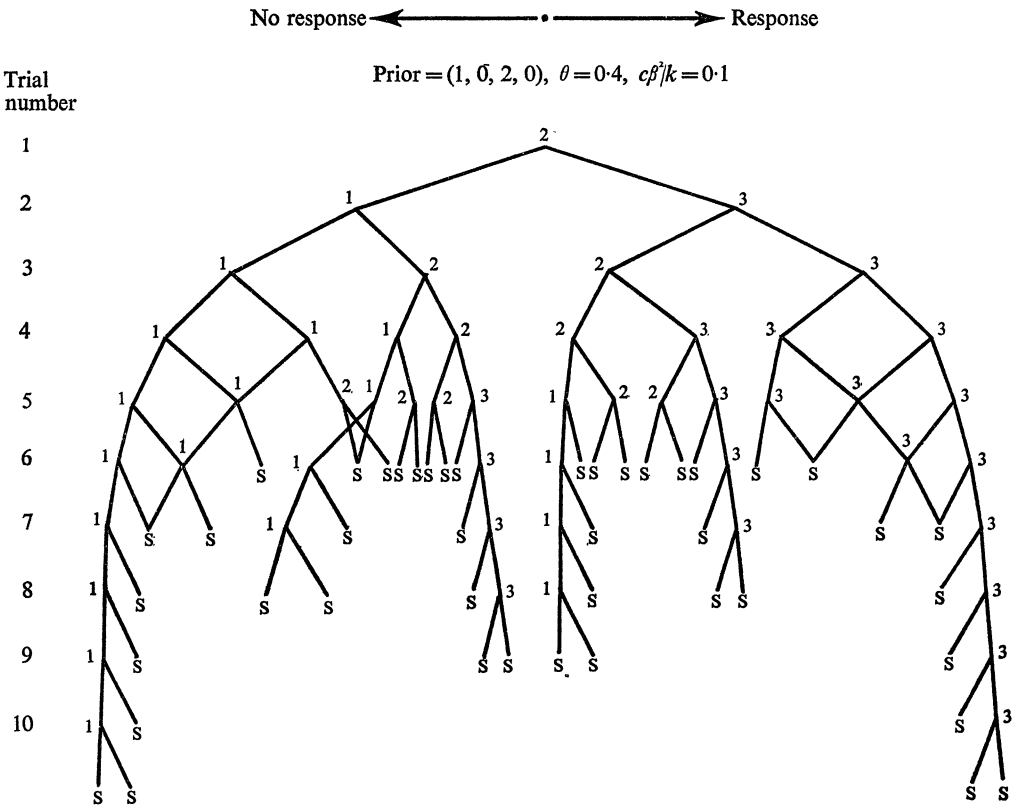

Fig. 4. Optimal sequential scheme using three dose levels. The digit at each branch point indicates the dose level to be used for the next trial. If a response is observed the branch to the right is taken, if a non-response then the branch to the left.

In the very simplified cases considered here, we obviously have to modify the up and down method somewhat. For example, with two dose levels $x_2 < x_1$, a response to a trial at level x_1 will cause the next trial to be at x_2, while a non-response will lead to a further trial at x_1. Although this is degrading the method seriously, it must be remembered that the unknown general Bayesian scheme has been equally degraded.

The efficiency results for the one, two and three level cases are shown in Tables 1, 2 and 3. The expected losses are all given in units of k/β^2. When the Bayesian expected sample size falls nearly equidistant between two integers, the up and down schemes corresponding to both were investigated but only one is given here.

6. Discussion

Probably the most surprising general result evident from Tables 1, 2 and 3 is the high overall efficiency of the up and down method. Only in the one level situation do the sample sizes get big enough for asymptotic tendencies to appear, and here it is remarkable how

Table 1. *Comparison of Bayesian and up and down methods using one dose level. Prior* (1, 2)

	Bayesian		Up and down		
$c\beta^2/k$	Expected sample size	Expected loss	Sample size	Expected loss	Efficiency
0·5	1	2·790	1	2·790	100
0·2	3·5	2·059	4	2·076	99·20
0·1	6·09	1·581	6	1·607	98·35
0·05	9·83	1·186	10	1·223	96·95
0·02	17·78	0·792	18	0·836	94·74
0·01	26·66	0·577	27	0·621	92·91
0·005	39·40	0·417	39	0·458	91·01
0·002	64·71	0·269	65	0·305	88·29
0·001	93·44	0·192	93	0·223	86·19
0·0005	134·28	0·137	134	0·163	84·14
0·0002	215·55	0·087	216	0·107	81·47
0·0001	307·21	0·062	307	0·078	79·45

Expected losses are throughout expressed in units of k/β^2.

Table 2. *Comparison of Bayesian and up and down methods using two dose levels*

			Bayesian		Up and down		
Prior distribution	$c\beta^2/k$	θ	Expected sample size	Expected loss	Sample size	Expected loss	Efficiency (%)
(1, 0, 2)	0·5	0·4	1·500	2·746	2	2·946	93·21
		0·1	2·137	2·730	2	2·745	99·45
		0·01	1·000	2·790	1	2·790	100·00
	0·4	0·4	2·252	2·538	2	2·746	92·43
		0·1	2·202	2·511	2	2·745	91·50
		0·01	2·376	2·585	2	2·668	96·89
	0·2	0·4	4·196	1·785	4	2·009	88·87
		0·1	3·570	1·952	4	1·999	97·64
	0·1	0·4	6·230	1·499	6	1·537	97·51
		0·1	6·441	1·471	6	1·525	96·46
(1, 1, 1)	0·5	0·4	2·000	2·766	2	2·766	100·00
		0·1	2·283	2·849	2	2·861	99·56
		0·01	2·201	3·367	2	3·427	98·25
	0·4	0·4	2·324	2·565	2	2·566	99·95
		0·1	2·465	2·615	2	2·661	98·26
		0·01	3·106	3·093	3	3·126	98·95
	0·2	0·4	3·724	1·986	4	2·012	98·73
		0·1	4·221	1·986	4	2·025	98·10
	0·1	0·4	6·787	1·449	7	1·541	94·07
		0·1	6·528	1·466	7	1·466	100·05

little is lost by using fixed sample sizes. When $c\beta^2/k = 0{\cdot}005$, for example, it can be seen from Fig. 1 that the optimal sample size can take values between 27 and 200, with a prior expectation of 39·40. Using exactly 39 trials, however, only increases the prior expectation of loss by $(8{\cdot}99/91{\cdot}01) \times 100 = 9{\cdot}88\%$.

In the two and three level situations, the very slight loss of efficiency of the up and down method will usually be more than compensated for by its greater simplicity of use. It should be remembered, however, that choice of an inappropriate initial dose level or stopping rule could well lead to expected losses much greater than those found here.

Table 3. *Comparison of Bayesian and up and down methods using three dose levels*

		$\theta = 0{\cdot}4$				
		Bayesian		Up and down		
Prior distribution	$c\beta^2/k$	Expected sample size	Expected loss	Sample size	Expected loss	Efficiency (%)
(1, 2, 0, 0)	0·5	1·557	2·590	2	2·746	94·34
	0·4	2·265	2·486	2	2·546	97·64
	0·2	3·849	1·940	4	1·977	98·11
	0·1	6·213	1·461	6	1·613	90·55
(1, 0, 2, 0)	0·5	2·000	2·702	2	2·702	100
	0·4	2·505	2·487	3	2·556	97·27
	0·2	3·817	1·902	4	1·994	95·43
	0·1	6·036	1·415	6	1·454	97·34
(1, 1, 1, 0)	0·5	2·000	2·737	2	2·737	100
	0·4	2·231	2·515	2	2·537	99·12
	0·2	3·724	1·931	4	2·132	90·60
	0·1	6·215	1·438	6	1·479	97·26

Since in most practical cases the up and down method settles down quickly to an oscillation over those few dose levels close to γ, it is plausible to speculate that the high efficiency found here will hold generally for, at least, medium and large sample sizes, when more than three dose levels are available. The only situation in which the Bayesian scheme is likely to be outstandingly better is when inaccurate prior ideas about γ cause the initial dose level to be far away from γ. The up and down method would then have to adjust the dose level one step at a time while the Bayesian method should permit much larger jumps and hence a more rapid approach to γ.

I am grateful to Professor D. V. Lindley for several discussions on this problem, and to a referee for helpful comments on an earlier draft of this paper.

REFERENCES

BROWNLEE, K. A., HODGES, J. L. & ROSENBLATT, M. (1953). The up and down method with small samples. *J. Am. Statist. Ass.* **48**, 262–77.

CHUNG, K. L. (1954). On a stochastic approximation method. *Ann. Math. Statist.* **25**, 463–83.

COCHRAN, W. G. & DAVIS, M. (1963). Sequential experiments for estimating the median lethal dose. *Le plan d'expériences*, pp. 181–94. Centre Nationale de la Recherche Scientifique, Paris.

COCHRAN, W. G. & DAVIS, M. (1964). Stochastic approximation to the median effective dose in bioassay. *Stochastic models in Medicine and Biology*, pp. 281–300. Madison: University of Wisconsin Press.

COCHRAN, W. G. & DAVIS, M. (1965). The Robbins–Monro method for estimating the median lethal dose. *J. R. Statist. Soc.* B, **27**, 28–45.

DIXON, W. J. & MOOD, A. M. (1948). A method for obtaining and analysing sensitivity data. *J. Am. Statist. Ass.* **43**, 109–26.

FERGUSON, T. S. (1967). *Mathematical Statistics*. New York: Academic Press.

KESTEN, H. (1958). Accelerated stochastic approximation. *Ann. Math. Statist.* **29**, 41–9.

LINDLEY, D. V. & BARNETT, B. N. (1965). Sequential sampling: two decision problems with linear losses for binomial and normal random variables. *Biometrika* **52**, 507–32.

MARKS, B. L. (1962). Some optimal sequential schemes for estimating the mean of a cumulative normal quantal response curve. *J. R. Statist. Soc.* B **24**, 393–400.

ROBBINS, H. & MONRO, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22**, 400–7.

SACKS, J. (1958). Asymptotic distribution of stochastic approximation procedures. *Ann. Math. Statist.* **29**, 373–405.

WETHERILL, G. B. (1963). Sequential estimation of quantal response curves. *J. R. Statist. Soc.* B, **25**, 1–48.

WETHERILL, G. B., CHEN, H. & VASUDEVA, R. B. (1966). Sequential estimation of quantal response curves: a new method of estimation. *Biometrika* **53**, 439–54.

WOLFOWITZ, J. (1952). On the stochastic approximation method of Robbins and Monro. *Ann. Math. Statist.* **23**, 457–61.

[*Received July* 1969. *Revised October* 1969]