SEQUENTIAL DESIGN OF EXPERIMENTS

By Herman Chernoff[1]

Stanford University

1. Introduction. Considerable scientific research is characterized as follows. The scientist is interested in studying a phenomenon. At first he is quite ignorant and his initial experiments are preliminary and tentative. As he gathers relevant data, he becomes more definite in his impression of the underlying theory. This more definite impression is used to construct more informative experiments. Finally after a certain point he is satisfied that his evidence is sufficient to allow him to announce certain conclusions and he does so.

While this sequential searching for relevant and informative experiments is common, very little statistical theory has been directed in this direction. The general problem may reasonably be called that of sequential design of experiments. A truncated variation of this problem called the two-armed bandit problem has attracted some attention (see [1] and [5]). Up to now an optimal solution for the two-armed bandit problem has not been attained. The failure to solve the two-armed bandit problem and certain obvious associated results indicate strongly that while optimal strategies are difficult to characterize, *asymptotically* optimal results should be easily available. Here the term asymptotic refers to large samples. For the sequential design problems, large samples and small cost of experimentation are roughly equivalent.

In this paper we present a procedure for the sequential design of experiments where the problem is one of testing a hypothesis. Formally, we assume that there are two possible actions (terminal decisions) and a class of available experiments. After each observation, the statistician decides on whether to continue experimentation or not. If he decides to continue, he must select one of the available experiments. If he decides to stop he must select one of the two terminal actions.

For the special case where there are only a finite numer of states of nature and a finite number of available experiments this procedure will be shown to be "asymptotically optimal" as the cost of sampling approaches zero. The procedure can be partially described by saying that at each stage the experimenter acts as though he is almost convinced that $\hat{\theta}$, the current maximum likelihood estimate of the state of nature, is actually equal to or very close to the true state of nature.

In problems were the cost of sampling is not small, this procedure may leave something to be desired. More specifically, until enough data are accumulated, the procedure may suggest very poor experiments because it does not sufficiently distinguish between the cases where $\hat{\theta}$ is a poor estimate and where $\hat{\theta}$ is a good

Received July 5, 1958.

[1] This work was sponsored by the Office of Naval Research under Contract N6 onr-25140 (NR-342-022).

estimate. For small cost of experimentation initial bungling is relatively unimportant. It is hoped and expected that with minor modifications the asymptotically optimal procedure studied here can be adapted to problems with relatively large cost of experimentation.

The procedures studied make extensive use of the Kullback-Leibler information numbers (see [2] and [4]).

2. Preliminaries involving two simple hypotheses. The procedure presented in this paper may be motivated by an asymptotic study of the classical problem of sequentially testing a simple hypothesis vs. a simple alternative with only one available experiment. Suppose $H_0: \theta = \theta_0$ and $H_1: \theta = \theta_1$ are two simple hypotheses, and the experiment yields a random variable x whose density is $f_i(x)$ under H_i, $i = 0, 1$. The Bayes strategies are the Wald sequential likelihood-ratio tests. These are characterized by two numbers A and B and consist of reacting to the first n observations $x_1, x_2, \cdots, x_n$

by rejecting H_0 if $\quad S_n \geqq A,$
accepting H_0 if $\quad S_n \leqq B,$

and continuing sampling as long as $B < S_n < A$ where

$$S_n = \sum_{i=1}^{n} \log [f_1(x_i)/f_0(x_i)]. \tag{2.1}$$

The appropriate numbers A and B are determined by the *a priori* probability w of H_1, and the costs. These are the cost c per observation (which is assumed fixed), the loss r_0 due to rejecting H_0 when it is true and the loss r_1 due to accepting H_0 when it is false. The risks corresponding to a sequential strategy are given by

$$\begin{aligned} R_0 &= r_0\alpha + c\mathcal{E}(N \mid H_0) \\ R_1 &= r_1\beta + c\mathcal{E}(N \mid H_1) \end{aligned} \tag{2.2}$$

where α and β are the two probabilities of error, and N is the possibly random sample size. Of course A and B are determined so as to minimize

$$(1 - w)R_0 + wR_1 .$$

Suppose that c approaches zero. Then A and $-B$ are large and Wald's approximations [6] give

$$\begin{aligned} &\alpha \approx e^{-A}, \qquad \beta \approx e^{B}, \\ &\mathcal{E}(N \mid H_0) \approx -B/I_0, \quad \text{and} \quad \mathcal{E}(N \mid H_1) \approx A/I_1 \end{aligned} \tag{2.3}$$

where I_0 and I_1 are the Kullback-Leibler information numbers given by

$$I_0 = \int \log [f_0(x)/f_1(x)] f_0(x)\, dx \tag{2.4}$$

and

$$I_1 = \int \log [f_1(x)/f_0(x)] f_1(x)\, dx \tag{2.5}$$

and are assumed to exist finite and positive. Minimizing the approximation to $(1 - w)R_0 + wR_1$ we find that

$$A \approx -\log c + \log[I_1 r_0(1 - w)/w] \approx -\log c, \tag{2.6}$$

$$B \approx \log c + \log[(1 - w)/I_0 r_1 w] \approx \log c, \tag{2.7}$$

$$\alpha \approx wc/I_1 r_0(1 - w), \qquad \beta \approx c(1 - w)/I_0 r_1 w,$$

$$\mathcal{E}(N \mid H_0) \approx -\log c/I_0, \qquad \mathcal{E}(N \mid H_1) \approx -\log c/I_1,$$

$$R_0 \approx -c \log c/I_0, \quad \text{and} \quad R_1 \approx -c \log c/I_1. \tag{2.8}$$

Remarks:

1. These results can be verified more rigorously by using Wald's bounds on his approximations when they apply. Our later results will generalize these approximations of R_0 and R_1.

2. The risk corresponding to the optimum strategy is mainly the cost of experimentation.

3. The optimum strategy and its risks depend mainly on c, I_0 and I_1 and are relatively insensitive to the costs r_0 and r_1 of making the wrong decision and to the *a priori* probability w. Note that doubling r_0 and r_1 is equivalent to cutting c in half as far as the strategy is concerned. The consequent change in $\log c$ which determines A and B is relatively small. That is, $\log c$ is changed to $\log c - \log 2$ while $\log 2$ is small compared to $\log c$.

Suppose that the experimenter is given a choice of one of two experiments E_1 and E_2 but that the one chosen must be used exclusively throughout the sequential testing problem. We designate the information numbers by $I_0(E_1)$, $I_1(E_1)$, $I_0(E_2)$ and $I_1(E_2)$. If c is small and $I_0(E_1) > I_0(E_2)$ and $I_1(E_1) > I_1(E_2)$, then it makes sense to select E_1. If, on the other hand, $I_0(E_1) > I_0(E_2)$ and $I_1(E_1) < I_1(E_2)$, then E_1 would be preferable if H_0 were "true" and E_2 would be preferable if H_1 were. Since the true state of nature is not known, there is no clear cut reason to prefer E_1 to E_2 without resorting to the *a priori* probabilities of H_0 and H_1.

The above rather artificial problem illuminates the more natural one where, after each decision to continue experimentation, one can choose between E_1 and E_2. If the cost of sampling is very small, it may pay to continue sampling even though we are almost convinced about which is the true state of nature. Thus even if we feel quite sure it is H_0, we may still be willing to experiment further, and furthermore it would make sense to select E_1 or E_2 by comparing $I_0(E_1)$ with $I_0(E_2)$.

To be more formal we may select E_1 or E_2 by comparing $I_0(E_1)$ and $I_0(E_2)$ if the maximum likelihood estimate of θ based on the previous observations is θ_0 and by comparing $I_1(E_1)$ and $I_1(E_2)$ if the maximum likelihood estimate is θ_1. Such a procedure may be short of optimal. On the other hand if c is very small, the most damage that could occur would be due to the *nonoptimal* choice of experiment for the first few of the many observations which are expected.

The stopping rule for the single experiment case may be naturally interpreted in terms of *a posteriori* probability as follows: there are two numbers of the order of magnitude of c. Stop experimenting if the *a posteriori* probability of H_0 goes below the first number or if the *a posteriori* probability of H_1 goes below the second number. The expected sample sizes are relatively insensitive to variations in the stopping limits. More specifically, if the *a posteriori* probability limits are any numbers of the order of magnitude of c, α and β are of the order of magnitude of c and $E(N \mid H_0) \approx -\log c/I_0$ and $E(N \mid H_1) \approx -\log c/I_1$.

In view of the standard derivations of the Bayes procedures for the one experiment sequential testing problem, it seems natural to stop when the *a posteriori* probability of H_0 or H_1 go below numbers of the order of magnitude of c. An example of such a stopping rule is obtained by selecting A and B equal to $-\log c$ and $\log c$ respectively. Then if $E^{(i)}$ is the experiment selected on the ith trial and x_i is the outcome let $z_i = \log [f_1(x_i, E^{(i)})/f_0(x_i, E^{(i)})]$ where $f_1(x, E^{(i)})$ and $f_0(x, E^{(i)})$ are the densities of the outcome of $E^{(i)}$. Finally, after the nth experiment, continue sampling only if $S_n = \sum_{i=1}^{n} z_i$ lies between B and A.

3. A special problem involving composite hypotheses. Comparing two probabilities. In the preceding section a method was proposed for the sequential design problem for testing a simple hypothesis vs. a simple alternative. The situation becomes more complicated when the hypotheses are composite. To motivate our procedures for this more complex problem, we shall discuss heuristically the special problem of comparing two probabilities. This problem may be regarded as a prototype of the general sequential design problem for testing hypotheses. We shall devote our main attention to the design aspect and leave the stopping rule in a relatively unrefined state.

It is desired to compare the efficacy of two drugs. The experiments E_1 and E_2 consist of using the first and second drugs respectively. The outcome of these experiments are success or failure, success having probabilities p_1 and p_2 in the two experiments. The two hypotheses are $H_1: p_1 > p_2$ and $H_2: p_1 \leqq p_2$.

After n observations consisting of n_1 trials of drug 1 and n_2 trials of drug 2, which led to m_1 and m_2 successes respectively, the maximum-likelihood estimate of $\theta = (p_1, p_2)$ is given by $\hat{\theta}_n = (\hat{p}_{1n}, \hat{p}_{2n}) = (m_1/n_1, m_2/n_2)$. We shall select our next experiment according to the following idea. Consider the experiment which would be appropriate if we believed that θ were $\hat{\theta}_n$ and we were testing the hypothesis $\theta = \hat{\theta}_n$ vs. the simple alternative $\theta = \tilde{\theta}_n$ which is the "nearest" parameter point under the "alternative hypothesis." To be more specific suppose $m_1/n_1 > m_2/n_2$. Then $\hat{\theta}_n$ is an element of the set of θ for which H_1 is true. The "nearest" element under H_2 is not clearly defined. For the present let us define it as the maximum-likelihood estimate under H_2. Then

$$\tilde{\theta}_n = \left(\frac{m_1 + m_2}{n_1 + n_2}, \frac{m_1 + m_2}{n_1 + n_2}\right)$$

$$= \left[\left(\frac{n_1}{n_1 + n_2}\right)\hat{p}_1 + \left(\frac{n_2}{n_1 + n_2}\right)\hat{p}_2, \left(\frac{n_1}{n_1 + n_2}\right)\hat{p}_1 + \left(\frac{n_2}{n_1 + n_2}\right)\hat{p}_2\right].$$

Note that $\tilde{\theta}_n$ is a weighted average of $(\hat{p}_1, \hat{p}_1)$ and $(\hat{p}_2, \hat{p}_2)$ where the weights are proportional to the frequencies of E_1 and E_2. If we were testing $\theta = \hat{\theta}_n$ vs. $\theta = \tilde{\theta}_n$ and strongly believed in $\theta = \hat{\theta}_n$, we would *select the experiment E for which*

$$(3.1) \qquad I(\hat{\theta}_n, \tilde{\theta}_n, E) = \sum_x \log [f(x, \hat{\theta}_n, E)/f(x, \tilde{\theta}_n, E)]\, f(x, \hat{\theta}_n, E)$$

is as large as possible where $f(x, \theta, E)$ is the probability of the data for experiment E when θ is the value of the parameter.[2] Thus

$$(3.2) \qquad I[(p_1, p_2), (p_1^*, p_2^*), E_1] = p_1 \log [p_1/p_1^*] + (1 - p_1) \log [(1 - p_1)/(1 - p_1^*)]$$

$$(3.3) \qquad I[(p_1, p_2), (p_1^*, p_2^*), E_2] = p_2 \log [p_2/p_2^*] + (1 - p_2) \log [(1 - p_2)/(1 - p_2^*)].$$

It is clear from these expressions and from intuitive considerations that if p_1^* is close to p_1, E_1 is relatively uninformative and if p_2^* is close to p_2, E_2 is relatively uninformative. Thus if n_1 is much larger than n_2, $\hat{\theta}_n$ is close to $(\hat{p}_1, \hat{p}_1)$ and

$$I(\hat{\theta}_n, \tilde{\theta}_n, E_1) < I(\hat{\theta}_n, \tilde{\theta}_n, E_2)$$

and E_2 is called for. Similarly if n_1 is much smaller than n_2, E_1 is called for. For a specified $\hat{\theta}_n$, there is a unique proportion $\lambda(\hat{\theta}_n)$ such that the two informations are equal if $n_2/(n_1 + n_2) = \lambda(\hat{\theta}_n)$. If $n_2/(n_1 + n_2)$ exceeds $\lambda(\hat{\theta}_n)$, E_1 is called for.

In general, the set of (p_1^*, p_2^*) for which

$$I[(p_1, p_2), (p_1^*, p_2^*), E_1)] = I[(p_1, p_2), (p_1^*, p_2^*), E_2]$$

is easy to characterize. See Fig. 1. It seems clear that after many observations $\hat{\theta}_n$ will be close to $\theta = (p_1, p_2)$ and $\tilde{\theta}_n$ will be close to that point $\theta^* = \theta^*(\theta) = (p_1^*, p_2^*)$ for which $p_1^* = p_2^*$ and $I(\theta, \theta^*, E_1) = I(\theta, \theta^*, E_2)$. Furthermore the proportion of times that E_1 is applied in the long run is determined by the relation of θ^* to θ. That is to say if $\theta = (p_1, p_2)$ and $\theta^* = [(1 - \lambda)p_1 + \lambda p_2, (1 - \lambda)p_1 + \lambda p_2]$ then $n_2/(n_1 + n_2)$ will tend to be close to $\lambda = \lambda(\theta)$.

The point θ^* and the ratio λ can also be interpreted from another point of view. Essentially θ^* is that point under the alternative hypothesis for which $\max_E I(\theta, \theta^*, E)$ is minimized. (In a sense this property can also be interpreted to say that θ^* is the "nearest" point to θ under the alternative hypotheses.) At θ^*, it doesn't matter which experiment is selected. On the other hand if we regard $I(\theta, \varphi, E)$ as the payoff matrix of a game and an experiment were to be chosen to maximize $\min_{\varphi \in a} I(\theta, \varphi, E)$ [a is the set corresponding to the alternative hypothesis], the randomized maximin strategy would give E_1 and E_2 weights $1 - \lambda$ and λ respectively. Thus θ^* and λ correspond to the solutions of a two person zero sum game with payoff $I(\theta, \varphi, E)$ where one player (say nature)

[2] We work against the "nearest" alternative under the intuitive assumption that this is the alternative which will make our risk large and which we must guard against.

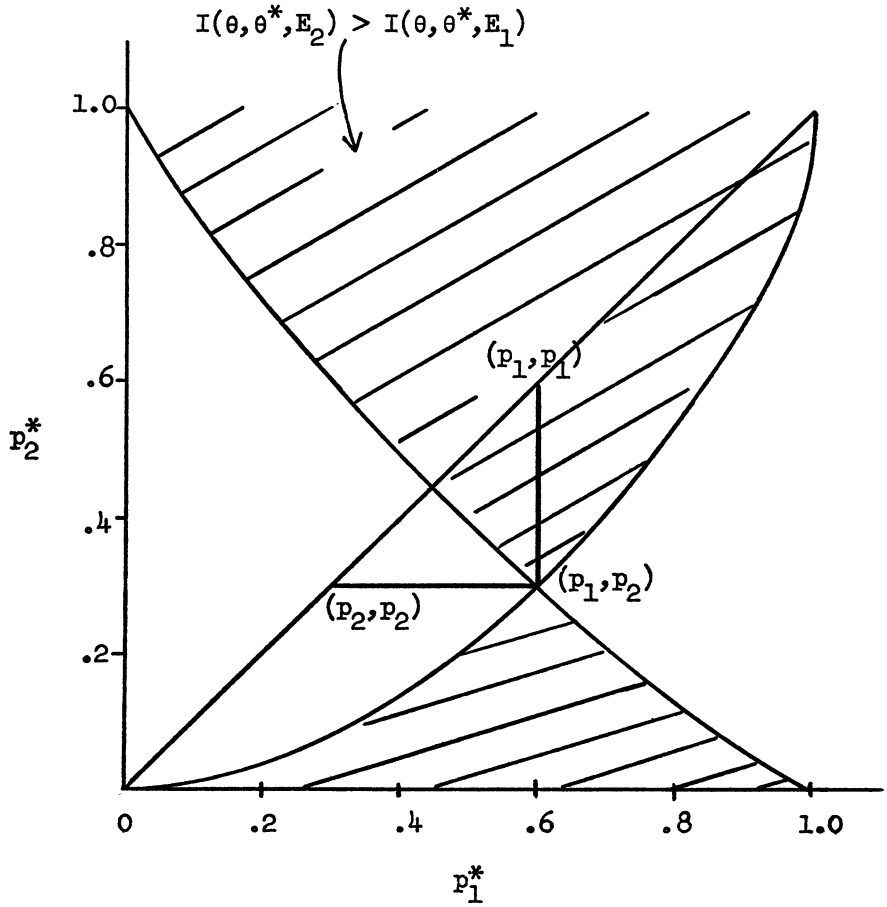

FIGURE 1

The set of $\theta^* = (p_1^*, p_2^*)$ *for which* E_2 *is preferred to* E_1, *i.e.,*

$$I(\theta, \theta^*, E_2) > I(\theta, \theta^*, E_1)$$

for a specified $\theta = (p_1, p_2)$ *in the drug testing problem*

selects φ to minimize I and the other player (the experimenter) selects E to maximize I.

It seems clear that the procedure recommended before will not be substantially affected, when c is small, if it is modified so that the n + 1st experiment is E_1 with probability $1 - \lambda_n$ and E_2 with probability λ_n where $1 - \lambda_n$ and λ_n correspond to the experimenter's maximin strategy based on the payoff matrix $I(\hat{\theta}_n, \varphi, E)$.

In Table 1, we tabulate as functions of θ,

a) $\theta^*(\theta)$, the minimax choice of nature,

TABLE 1

Tabulation[3] *of* $p^*(\theta)$, $\lambda(\theta)$, $I(\theta)$ *and* $e(\theta)$ *where* $\theta = (p_1, p_2)$ *and* $\theta^*(\theta) = [p^*(\theta), p^*(\theta)]$

p_1	p_2	.01	.05	.10	.20	.40
.05	p^*	.0276				
	λ	.560				
	I	.00329				
	e	.988				
.10	p^*	.0479	.0737			
	λ	.579	.526			
	I	.00997	.00200			
	e	.978	.995			
.20	p^*	.0889	.118	.148		
	λ	.585	.547	.520		
	I	.0258	.0120	.00435		
	e	.971	.992	.995		
.40	p^*	.171	.205	.239	.297	
	λ	.588	.557	.537	.515	
	I	.0637	.0429	.0277	.0105	
	e	.973	.988	.992	.999	
.60	p^*	.260	.297	.333	.394	.500
	λ	.576	.551	.534	.515	.500
	I	.111	.0855	.0648	.0375	.00874
	e	.980	.991	.995	.999	1.000
.80	p^*	.363	.400	.438	.500	
	λ	.553	.533	.517	.500	
	I	.174	.145	.120	.0837	
	e	.990	.996	.997	1.000	
.90	p^*	.425	.463	.500		
	λ	.538	.514	.500		
	I	.217	.187	.160		
	e	.998	.999	1.000		

b) $\lambda(\theta)$, the proportion of times E_2 is used in the experimenter's maximin strategy,

c) $I(\theta) = I(\theta, \theta^*(\theta), E_1) = I(\theta, \theta^*(\theta), E_2)$, the value of the game, and

d) $e(\theta) = \min_{\varphi \epsilon a} [I(\theta, \varphi, E_1) + I(\theta, \varphi, E_2)]/2I(\theta, \theta^*(\theta), E_1)$,

which represents a measure of the relative efficiency of using each drug half the time to the procedure advocated. Evidently there is no great loss of efficiency in using each drug half the time and this prototype example is mainly useful as an illustrative device.

[3] By symmetry we need only consider $p_1 < p_2$ and $p_1 + p_2 \leqq 1$.

4. Formal description of the general procedures. It is desired to test

$$H_1: \theta \in \omega_1$$

vs. the alternative $H_2: \theta \in \omega_2$. There is available a set of experiments $\{E\}$ each of which may be replicated. Let $f(x, \theta, E)$ be the density of the outcome x of experiment E with respect to a measure μ_E. Let

$$I(\theta_1, \theta_2, E) = \int \log \left[\frac{f(x, \theta_1, E)}{f(x, \theta_2, E)}\right] f(x, \theta_1, E)\, d\mu_E(x). \tag{4.1}$$

Designate the nth experiment selected by $E^{(n)}$. Although the choice of the $n+1$st experiment may depend on the past, once it is selected, its outcome is assumed independent of those of the preceding experiments. The maximum-likelihood estimate of θ based on the first n observations is designated by $\hat{\theta}_n$. If $\theta \in \omega_1$, we call H_1 the hypothesis of θ, H_2 the hypothesis alternative to θ, and $a(\theta) = \omega_2$ the set alternative to θ. If $\theta \in \omega_2$, the hypothesis of θ is H_2, the alternative hypothesis is H_1 and the alternative set is $a(\theta) = \omega_1$. Let $\tilde{\theta}_n$ be the maximum likelihood estimate of θ under the hypothesis alternative to $\hat{\theta}_n$. If ω_1, ω_2, or the union $\omega_1 \cup \omega_2$ are not "closed" we may have some difficulty due to the nonexistence of these estimates. We shall assume throughout that suitably closing ω_1 and ω_2 and, if necessary, taking $\hat{\theta}_n$ and $\tilde{\theta}_n$ on the boundary of ω_1 and ω_2 eliminates this difficulty. In particular in the drug testing problem $\tilde{\theta}_n$ lies on the boundary separating ω_1 and ω_2.

Let $E(\theta, \varphi)$ be any experiment which maximizes $I(\theta, \varphi, E)$. We assume that such information maximizing experiments exist. Let us regard $I(\theta, \varphi, E)$ as the payoff matrix of a two-person zero-sum game where one player (seeking to minimize I) selects φ among the elements of the closure of $a(\theta)$, the set alternative to θ, and the other player (seeking to maximize I) selects E as an element of $\{E\}$, the set of available strategies. Now suppose that the experiment E is selected by some random mechanism corresponding to a probability measure λ on $\{E\}$. Then the corresponding information is $\int I(\theta, \varphi, E)\, d\lambda(E)$. Since experiments may be selected in such a fashion we extend the space of available experiments to include this convex set of all randomized experiments. This set is equivalent to the class of all randomized strategies of the second player (experimenter) in the original game. Thus if the original game had randomized or pure solutions, the second player in the extended game has a (not necessarily unique) maximin strategy $E(\theta)$ which is a pure or randomized experiment. The value of the game is given by

$$I(\theta) = \inf_{\varphi \in a(\theta)} I(\theta, \varphi, E(\theta)). \tag{4.2}$$

Let $E^{(n)}$ represent the experiment used for the nth trial, x_n the corresponding outcome,

$$z_n(\theta, \varphi) = \log [f(x_n, \theta, E^{(n)})/f(x_n, \varphi, E^{(n)})], \tag{4.3}$$

$$S_n(\theta, \varphi) = \sum_{i=1}^{n} z_i(\theta, \varphi), \tag{4.4}$$

and

$$S_n = \sum_{i=1}^{n} z_i(\hat{\theta}_n, \tilde{\theta}_n). \tag{4.5}$$

We define our procedure A as follows. *Stop sampling at the nth observation and select the hypothesis of* $\hat{\theta}_n$ *if* $S_n > -\log c$. *If sampling is to be continued let* $E^{(n+1)} = E(\hat{\theta}_n)$ which is to be defined in some unique and measurable way consistent with the above definition of $E(\theta)$.

The stopping rule described above is rather unrefined. It does not require much imagination to see how to go about refining it. On the other hand such refinements will not be necessary for the asymptotic results we want and so we shall not study them here.

5. Asymptotic characteristics of procedure A. In this section we shall evaluate the asymptotic characteristics of the procedure A of Section 4 for the case where there are only a finite number of states of nature and a finite number of available (pure) experiments. We shall show that for this procedure the probability of making the wrong decision is $O(c)$ and the expected sample size is asymptotically no larger than $-\log c/I(\theta)$. In the next section we shall show that this is as well as can be done.

The arguments will involve the fact that as $c \to 0$, the required sample size gets large. For large samples, $\hat{\theta}_n$ tends to be equal to θ for all but the first "few" observations. Then, $E^{(n)} = E(\theta)$ and for $\varphi \in a(\theta)$, $\mathcal{E}\{z_n(\theta, \varphi)\} = I[\theta, \varphi, E(\theta)] \geqq I(\theta)$ and the sample size required for S_n to reach $-\log c$ is approximately $-\log c/I(\theta)$.

For the results in the rest of this paper we make the following assumptions. *There are s possible states of nature which are divided into the two disjoint sets* ω_1 *and* ω_2. *The set of available (pure) experiments is* $\{E_1, E_2, \cdots, E_k\}$. *The losses due to making the wrong decision are assumed to be positive. That is,* $r(\theta, i)$, *which is equal to the loss due to selecting* H_i *when* θ *is the state of nature, satisfies*

$$\begin{aligned} r(\theta, i) &= 0 \quad \text{if} \quad \theta \in \omega_i, \qquad i = 1, 2 \\ r(\theta, i) &> 0 \quad \text{if} \quad \theta \notin \omega_i, \qquad i = 1, 2. \end{aligned} \tag{5.1}$$

If the experiment E yields outcome x and θ *and* φ *are distinct, then*

$$z(\theta, \varphi, E) = \log \frac{f(x, \theta, E)}{f(x, \varphi, E)} \tag{5.2}$$

has finite variance[4] *and*

$$I(\theta, \varphi, E) = \int \log \left[\frac{f(x, \theta, E)}{f(x, \varphi, E)}\right] f(x, \theta, E)\, d\mu_E(x) > 0. \tag{5.3}$$

[4] If $z(\theta, \varphi, E_i)$ has finite variance for each of the finite number of pure experiments, it follows easily that the variance of $z(\theta, \varphi, E)$ is bounded for all randomized experiments. Similarly $I(\theta, \varphi, E)$ is bounded away from zero and infinity.

Hereafter we represent the true state of nature by θ_0 which we assume to be in ω_1. Unless clearly specified otherwise, all probabilities and expectations refer to θ_0.

LEMMA 1: *If the stopping rule is disregarded and sampling is continued according to any measurable*[5] *procedure,*

$$\hat{\theta}_n \to \theta_0 \quad \text{w.p.1.} \tag{5.4}$$

In fact there exist K *and* $b > 0$ *such that*

$$P\{T > n\} \leqq Ke^{-bn} \tag{5.5}$$

where T *is defined as the smallest integer such that* $\hat{\theta}_n = \theta_0$ *for* $n \geqq T$.

PROOF: Assign *a priori* probability $1/s$ to each state of nature. Then $\hat{\theta}_n$ is the value of θ which maximizes the *a posteriori* probability $p_n(\theta)$ after the nth observation. Furthermore,

$$\log\left[\frac{p_n(\theta_0)}{p_n(\theta)}\right] = \sum_{i=1}^{n} \log\left[\frac{f(x_i, \theta_0, E^{(i)})}{f(x_i, \theta, E^{(i)})}\right] = \sum_{i=1}^{n} z_i(\theta_0, \theta) = S_n(\theta_0, \theta) \tag{5.6}$$

and $\hat{\theta}_n = \theta_0$ if $S_n(\theta_0, \theta) > 0$ for all $\theta \neq \theta_0$. It suffices to show that for each $\theta \neq \theta_0$, there is a number b such that $P\{S_n(\theta_0, \theta) \leqq 0\} \leqq e^{-bn}$. But

$$\begin{aligned} &P\{S_n(\theta_0, \theta) \leqq 0\}\mathcal{E}\{e^{tS_n(\theta_0,\theta)} \mid S_n(\theta_0, \theta) \leqq 0\} \leqq \mathcal{E}\{e^{tS_n(\theta_0,\theta)}\} \\ &P\{S_n(\theta_0, \theta) \leqq 0\} \leqq \mathcal{E}\{e^{tS_n(\theta_0,\theta)}\} \quad \text{for} \quad t \leqq 0. \end{aligned} \tag{5.7}$$

Now $\mathcal{E}\{[f(x, \theta_0, E)/f(x, \theta, E)]^t\}$ is the moment generating function of

$$z(\theta_0, \theta, E) = \log [f(x, \theta_0, E)/f(x, \theta, E)]$$

and is equal to one for $t = -1$ and $t = 0$ and, by convexity, is less than one for $-1 < t < 0$. Thus for a randomized experiment E where E_i is selected with probability p_i,

$$\mathcal{E}\{e^{-z(\theta_0,\theta,E)/2}\} = \sum_{i=1}^{k} p_i \mathcal{E}\{e^{-z(\theta_0,\theta,E_i)/2}\}$$

is bounded below one, and there is a number $b > 0$ such that

$$\max_E \mathcal{E}\{e^{-z(\theta_0,\theta,E)/2}\} = e^{-b} < 1.$$

Thus

$$\mathcal{E}\{e^{-z_i(\theta,\theta)/2} \mid x_1, x_2, \cdots, x_{i-1}\} \leqq e^{-b}$$

and

$$P\{S_n(\theta_0, \theta) \leqq 0\} \leqq \mathcal{E}\{e^{-S_n(\theta,\theta)/2}\} \leqq e^{-bn}.$$

LEMMA 2: *For procedure* A, *the expected sample size satisfies*

$$\mathcal{E}(N) \leqq -[1 + o(1)] \log c/I(\theta_0). \tag{5.8}$$

[5] A procedure is considered measurable, if at the nth stage, the experiment selected is a measurable function of the data $x_1, x_2, \cdots, x_n$.

PROOF: We wish to show that, given any $\epsilon > 0$, there is a $c^* = c^*(\epsilon)$ such that

$$\mathcal{E}(N) \leqq -(1 + \epsilon) \log c/I(\theta_0) \quad \text{for} \quad c < c^*.$$

It is obvious that $N \leqq \max_{\varphi \varepsilon \omega_2} (N_\varphi, T)$ where N_φ is the smallest integer for which $\sum_{i=1}^{n} z_i(\theta_0, \varphi) > -\log c$ for all $n \geqq N_\varphi$. In view of Lemma 1, it suffices to show that for each $\varphi \varepsilon \omega_2$ and for each $\epsilon > 0$, there exist $K = K(\epsilon, \varphi)$ and $b = b(\epsilon, \varphi) > 0$ such that

$$P\{N_\varphi > n\} \leqq Ke^{-bn} \quad \text{for} \quad n > -(1 + \epsilon) \log c/I(\theta_0).$$

To show this it suffices to prove that for each $\varphi \varepsilon \omega_2$ and each $\epsilon > 0$, there exist K and $b > 0$ such that

$$(5.9) \qquad P\left\{\sum_{i=1}^{n} z_i(\theta_0, \varphi) < -\log c\right\} \leqq Ke^{-bn} \quad \text{for} \quad n > -(1 + \epsilon) \log c/I(\theta_0).$$

But

$$\sum_{i=1}^{n} z_i(\theta_0, \varphi) = \sum_{i=1}^{n} [z_i(\theta_0, \varphi) - I(\theta_0, \varphi, E^{(i)})] + \sum_{i=1}^{n} [I(\theta_0, \varphi, E^{(i)}) - I(\theta_0, \varphi, E(\theta_0))] + nI(\theta_0, \varphi, E(\theta_0)).$$

If $\epsilon_1 > 0$, $z(\theta_0, \varphi, E_i) - I(\theta_0, \varphi, E_i) + \epsilon_1$ has positive mean and finite moment generating function for $-1 \leqq t \leqq 0$ for each φ and pure experiment E_i. Hence the left-hand derivative of the moment generating function is positive at $t = 0$. Thus there is a $t^* = t^*(\epsilon) < 0$ and $b_1 = b_1(\epsilon) > 0$ such that

$$\mathcal{E}\{e^{t^*[z(\theta_0, \varphi, E_i) - I(\theta_0, \varphi, E_i) + \epsilon_1]}\} \leqq e^{-b_1}.$$

Consequently,

$$\mathcal{E}\{e^{t^*[z(\theta_0, \varphi, E) - I(\theta_0, \varphi, E) + \epsilon_1]}\} \leqq e^{-b_1}$$

for each φ and E. Then, as in our proof of Lemma 1,

$$\mathcal{E}\{e^{t^*\left[\sum_{i=1}^{n} z_i(\theta_0, \varphi) - I(\theta_0, \varphi, E^{(i)}) + \epsilon_1\right]}\} \leqq e^{-b_1 n}$$

and

$$(5.10) \qquad P\left\{\sum_{i=1}^{n} [z_i(\theta_0, \varphi) - I(\theta_0, \varphi, E^{(i)})] < -\epsilon_1 n\right\} \leqq e^{-b_1 n}.$$

Furthermore, it follows from the definition of T that

$$\left|\sum_{i=1}^{n} [I(\theta_0, \varphi, E^{(i)}) - I(\theta_0, \varphi, E(\theta_0))]\right| \leqq K_2 T$$

and hence, applying Lemma 1,

$$(5.11) \qquad P\left\{\sum_{i=1}^{n} [I(\theta_0, \varphi, E^{(i)}) - I(\theta_0, \varphi, E(\theta_0))] < -\epsilon_2 n\right\} \leqq K_3 e^{-b_3 n}.$$

Finally,

$$I(\theta_0, \varphi, E(\theta_0)) \geqq I(\theta_0) \qquad \text{for } \varphi \in \omega_2. \tag{5.12}$$

Then combining (5.10), (5.11), and (5.12) we obtain

$$P\left\{\sum_{i=1}^{n} z_i(\theta_0, \varphi) < n[I(\theta_0) - \epsilon_3]\right\} \leqq K_4 e^{-b_4 n}$$

from which the desired result (5.9) follows.

LEMMA 3: *For procedure A, the probability of error (rejecting H_1) is $\alpha = O(c)$.*

PROOF: On the set $A_{n\varphi}$ in the sample space for which we reject $H_1 : \theta \in \omega_1$ at the nth observation and for which $\hat{\theta}_n = \varphi \in \omega_2$

$$\sum_{i=1}^{n} z_i(\varphi, \theta_0) \geqq \sum_{i=1}^{n} z_i(\hat{\theta}_n, \tilde{\theta}_n) \geqq -\log c$$

and

$$\prod_{i=1}^{n} f(x_i, \theta_0, E^{(i)}) \leqq c \prod_{i=1}^{n} f(x_i, \varphi, E^{(i)}).$$

$$P\{A_{n\varphi}\} = \int_{A_{n\varphi}} \prod_{i=1}^{n} f(x_i, \theta_0, E^{(i)})\, d\mu_{E^{(1)}}(x_1) \cdots d\mu_{E^{(n)}}(x_n)$$

$$\leqq c \int_{A_{n\varphi}} \prod_{i=1}^{n} f(x_i, \varphi, E^{(i)})\, d\mu_{E^{(1)}}(x_1) \cdots d\mu_{E^{(n)}}(x_n).$$

The last integral is the probability of the set $A_{n\varphi}$ when φ is the state of nature. Thus

$$\alpha = \sum_{\varphi \in \omega_2} \sum_{n=1}^{\infty} P\{A_{n\varphi}\} \leqq \sum_{\varphi \in \omega_2} c < sc = O(c).$$

This proof is rather standard and obviously applies to any measurable procedure with the same stopping rule as procedure A.

Combining Lemmas 2 and 3 we have

THEOREM 1: *For procedure A, the risk function $R(\theta)$ satisfies*

$$R(\theta) \leqq -[1 + o(1)]c \log c/I(\theta) \tag{5.13}$$

for all θ.

6. Asymptotic optimality of procedure A. We shall state and prove Theorem 2 which together with Theorem 1 will establish the asymptotic optimality of procedure A in the sense described below.

THEOREM 2: *Any procedure for which $I(\theta) > 0$ and*

$$R(\theta) = O(-c \log c) \qquad \text{for all } \theta$$

satisfies

$$R(\theta) \geqq -[1 + o(1)]c \log c/I(\theta) \qquad \text{for all } \theta. \tag{6.1}$$

Combining Theorems 1 and 2 we see that for any procedure to do substantially better than A for any θ implies that its risk will be of a *greater order of magnitude* for some θ. In this sense procedure A is asymptotically optimal.

To prove Theorem 2 we shall use two lemmas. The first will show that for the probabilities of error to be small enough,

$$S(\theta_0, \varphi) = \sum_{i=1}^{N} z_i(\theta_0, \varphi) \tag{6.2}$$

must be sufficiently large for all φ in ω_2 with large probability. The second will show that when n is substantially smaller than $-\log c/I(\theta_0)$, it is unlikely that the sums $\sum_{i=1}^{n} z_i(\theta_0, \varphi)$ can be sufficiently large for all $\varphi \varepsilon \omega_2$.

LEMMA 4: *If* $\varphi \varepsilon \omega_2$, $P\{accept\ H_1 \mid \theta = \varphi\} = O(-c \log c)$, $P\{reject\ H_1\} = O(-c \log c)$, *and* $0 < \epsilon < 1$, *then*

$$P\{S(\theta_0, \varphi) < -(1 - \epsilon) \log c\} = O(-c^{\epsilon} \log c). \tag{6.3}$$

PROOF: There is a number K such that

$$-Kc \log c \geqq P\{\text{accept } H_1 \mid \theta = \varphi\} \geqq \sum_{n=1}^{\infty} \int_{A_n} f(x, \varphi)\, d\mu(x)$$

where $f(x, \varphi)$ is the density on the sample space when $\theta = \varphi$ and A_n is the subset of the sample space for which $S(\theta_0, \varphi) < -(1 - \epsilon) \log c$ and H_1 is accepted at the nth step.

$$-Kc \log c \geqq \sum_{n=1}^{\infty} \int_{A_n} [f(x, \varphi)/f(x, \theta_0)] f(x, \theta_0)\, d\mu(x)$$

$$= \sum_{n=1}^{\infty} \int_{A_n} e^{-S(\theta_0, \varphi)} f(x, \theta_0)\, d\mu(x) \geqq c^{1-\epsilon} \sum_{n=1}^{\infty} P\{A_n\}.$$

$$\sum_{n=1}^{\infty} P\{A_n\} = O(-c^{\epsilon} \log c).$$

$$P\{\text{reject } H_1\} = O(-c \log c).$$

$$P\{S(\theta_0, \varphi) < -(1 - \epsilon) \log c\} \leqq \sum_{n=1}^{\infty} P\{A_n\} + P\{\text{reject } H_1\} = O(-c^{\epsilon} \log c).$$

LEMMA 5: *If* $\epsilon > 0$,

$$P\left\{\max_{1 \leqq m \leqq n} \min_{\varphi \varepsilon \omega_2} \sum_{i=1}^{m} z_i(\theta_0, \varphi) \geqq n[I(\theta_0) + \epsilon]\right\} \to 0 \quad as \quad n \to \infty. \tag{6.4}$$

PROOF:

$$\sum_{i=1}^{m} z_i(\theta_0, \varphi) = \sum_{i=1}^{m} [z_i(\theta_0, \varphi) - I(\theta_0, \varphi, E^{(i)})] + \sum_{i=1}^{m} I(\theta_0, \varphi, E^{(i)})$$

$$= A_{1m} + A_{2m},$$

where

$$A_{1m} = \sum_{i=1}^{m} [z_i(\theta_0, \varphi) - I(\theta_0, \varphi, E^{(i)})]$$

is a martingale. Now $A_{2m} = \sum_{i=1}^{m} I(\theta_0, \varphi, E^{(i)})$ represents m times the payoff for the game where nature selects φ and the experimenter selects some mixture of his available strategies. Thus $\min_{\varphi \epsilon \omega_2} A_{2m} \leqq mI(\theta_0) \leqq nI(\theta_0)$. Thus

$$\min_{\varphi \epsilon \omega_2} (A_{1m} + A_{2m}) \geqq n[I(\theta_0) + \epsilon]$$

implies that $A_{1m} \geqq n\epsilon$ for some φ and

$$P\left\{\max_{1 \leqq m \leqq n} \min_{\varphi \epsilon \omega_2} \sum_{i=1}^{m} z_i(\theta_0, \varphi) \geqq n[I(\theta_0) + \epsilon]\right\} \leqq \sum_{\varphi \epsilon \omega_2} P\{\max_{1 \leqq m \leqq n} A_{1m} \geqq n\epsilon\}.$$

Since A_{1m} is a martingale with mean 0, we may apply Doob's extension of Kolmogoroff's extension of Tchebycheff's inequality ([3], p. 315), to obtain

$$P\{\max_{1 \leqq m \leqq n} A_{1m} > n\epsilon\} \leqq K/n\epsilon^2 \quad \text{for each} \quad \varphi \, \varepsilon \, \omega_2.$$

Lemma 5 follows.

Now we are in position to prove Theorem 2. Let

$$n_c = -(1 - \epsilon) \log c/[I(\theta_0) + \epsilon];$$

$$P\{N \leqq n_c\} \leqq P\{N \leqq n_c \quad \text{and} \quad S(\theta_0, \varphi) \geqq -(1 - \epsilon) \log c \quad \text{for all } \varphi \, \varepsilon \, \omega_2\}$$
$$+ P\{S(\theta_0, \varphi) < -(1 - \epsilon) \log c \quad \text{for some } \varphi \, \varepsilon \, \omega_2\}.$$

By Lemma 5, the first term on the right approaches zero. The condition $R(\theta) = O(-c \log c)$ for all θ permits us to apply Lemma 4 and the second term on the right approaches zero. Hence $\mathcal{E}(N) \geqq -[1 + o(1)] \log c/I(\theta_0)$. Theorem 2 follows.

7. Miscellaneous remarks.

1. The asymptotic optimality of procedure A may not be especially relevant for the initial stages of experimentation especially if the cost of sampling is high. At first it is desirable to apply experiments which are informative for a broad range of parameter values. Maximizing the Kullback-Leibler information number may give experiments which are efficient only when θ is close to the estimated value.

2. It is clear that the methods and results apply when the cost of sampling varies from experiment to experiment. Here, we are interested in selecting experiments which maximize information per unit cost.

3. The ideas employed in this paper seem equally valid and applicable to problems which involve selecting one of k mutually exclusive hypotheses.

4. A minor modification of the stopping rule would be to continue experimentation as long as $\sum_{i=1}^{n} z_i(\hat{\theta}_i, \tilde{\theta}_i) < -\log c$. This rule which involves

$$\sum_{i=1}^{n} z_i(\hat{\theta}_i, \tilde{\theta}_i)$$

instead of $\sum_{i=1}^{n} z_i(\hat{\theta}_n, \tilde{\theta}_n)$ may be computationally easier to deal with occasionally. It is not difficult to show that Lemma 2 applies for this stopping rule. The author has not proved that Lemma 3 also applies.

5. A modification of the experimentation rule is the following. Select $E^{(n+1)}$ so as to maximize[6] $I(\hat{\theta}_n, \tilde{\theta}_n, E)$. It is easy to see that Lemma 3 would apply for this or for any measurable experimentation rule. While it is expected that Lemma 2 would apply for some examples the example below seems to indicate that it should not apply in general for this modified experimentation rule.

Note that for large samples we will have

$$\sum_{i=1}^{n} z_i(\theta_0, \varphi) \approx n \sum_{i=1}^{k} \frac{m_{in}}{n} I(\theta_0, \varphi, E_i)$$

where m_{in} is the number of times E_i is applied in the first n experiments.

Assuming $\hat{\theta}_n = \theta_0 \, \varepsilon \, \omega_1$, $\tilde{\theta}_n$ is that value of $\varphi \, \varepsilon \, \omega_2$ which minimizes

$$\textstyle\sum_{i=1}^{n} z_i(\theta_0, \varphi).$$

Thus $\tilde{\theta}_n$ essentially minimizes $\sum_{i=1}^{k} (m_{in}/n) I(\theta_0, \varphi, E_i)$. The successive choices of $\tilde{\theta}_n$ and $E^{(n+1)}$ correspond to the following strategies of two players of a game. Player 1 sees what strategy repeated n times would have been most effective against the combination of the past choices of player 2. Player 2 selects $E^{(n+1)}$ as though player 1 would select that most effective strategy. If for this iterative choice we have

$$\min_{\varphi \varepsilon \omega_2} \left[\sum_{i=1}^{n} I(\theta_0, \varphi, E^{(i)}) \right] \Big/ n \geqq I(\theta_0) - o(1)$$

Lemma 2 should apply. The following example shows that we can not always obtain the above inequality if there are more than 3 available experiments. [In our prototype example, it is quite clear that no such difficulty will arise.] Let $I(\theta_0, \varphi, E)$ be given by the following table.

TABLE 2

$I(\theta_0, \varphi, E)$

	E_1	E_2	E_3
φ_1	10	1	6
φ_2	1	10	6

Then our iterative procedure will always lead to E_1 or E_2. In fact each will be used approximately half the time giving a limiting value of 5.5 for

$$\sum_{i=1}^{n} I(\theta_0, \varphi_j, E^{(i)})/n, \qquad j = 1, 2.$$

On the other hand $I(\theta_0) = 6$. Clearly this modified experimentation rule is not as dependable as the one we chose.

[6] This rule was essentially the first one suggested in our study of the prototype example of Section 3.

6. The asymptotic study of the problem of testing a simple hypothesis vs. a simple alternative suggests that it should be possible to refine the stopping rule for the composite problem. While the main term of the risk should not be affected the higher order terms could probably be improved. Such improvement may be quite important in the case where c is not very small. A refinement in the stopping rule would be relevant for problems of testing composite hypotheses even if the problems do not involve the choice of experiments.

7. In Equation (5.3) we require that $I(\theta, \varphi, E) > 0$. This condition, used in the proof of consistency in Lemma 1, is not satisfied in the drug testing problem. There, using drug 1 will give $I(\theta, \varphi, E_1) = 0$ if $\theta = (p_1, p_2)$ and $\varphi = (p_1, p_2^*)$. However, this condition can be relaxed, if procedure A is modified slightly to assure consistency. For example, let E be a specified mixture involving each of the pure experiments. If we use E instead of $E^{(n)}$ whenever n is a perfect square, we will have the desired consistency so long as there is an E_i for each θ and φ such that $I(\theta, \varphi, E_i) > 0$. Even this may be relaxed since it is not necessary to discriminate between θ and φ if they correspond to the same hypothesis. In fact, it suffices to have $I(\theta) > 0$ for each θ.

REFERENCES

[1] R. N. Bradt, S. M. Johnson, and S. Karlin, "On sequential designs for maximizing the sum of n observations," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 1060–1074.

[2] Herman Chernoff, "Large-sample theory; parametric case," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 1–22.

[3] J. L. Doob, *Stochastic Processes*, John Wiley and Sons, Inc., New York, 1953.

[4] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 79–86.

[5] H. E. Robbins, "Some aspects of the sequential design of experiments," *Bull. Amer. Math. Soc.*, Vol. 58, No. 5 (1952), pp. 527–535.

[6] Abraham Wald, *Sequential Analysis*, John Wiley and Sons, Inc., New York, 1947.