ASYMPTOTIC BEHAVIOR OF SEQUENTIAL DESIGN WITH COSTS OF EXPERIMENTS

By Kazutomo Kawamura

1. Introduction.

We shall consider the two kinds of experiments E_1 and E_2 which have two events "Success S" or "Feilure F". The probabilities of success or failure by the experiments E_1 and E_2 are given by

$$P\{S|E_1\}=p_1, \qquad P\{F|E_1\}=1-p_1$$

and

$$P\{S|E_2\}=p_2, \qquad P\{F|E_2\}=1-p_2$$

respectively, where we assume that $p_1 \neq p_2$.

Moreover, following to Kunisawa [4], we introduce the notion of costs of experiments, i.e., if we execute the experiment E_1, it costs c_1 ($c_1>0$), and if E_2, it costs c_2 ($c_2>0$).

The object of this paper is to discriminate the hypotheses $p_1>p_2$ or $p_1<p_2$. What a procedure, with which we repeat the experiments, is optimal, in order to maximize the information of discrimination per unit cost?

According to Chernoff [1] a procedure is given, which maximizes the information, when $c_1=c_2$.

In this paper we shall show the asymptotic behavior of the procedure which maximizes the information of discrimination per unit cost.

2. Notations and definitions.

Given Θ the two dimensional closed rectangular set $[0, 1]\otimes[0, 1]$, i.e., the set of elements (p_1, p_2) satisfying $0 \leqq p_1 \leqq 1$ and $0 \leqq p_2 \leqq 1$. And put

$$H_1=\{(p_1, p_2): p_1>p_2, (p_1, p_2)\in\Theta\},$$

$$H_2=\{(p_1, p_2): p_1<p_2, (p_1, p_2)\in\Theta\}$$

and

$$B_{12}=\{(p_1, p_2): p_1=p_2, (p_1, p_2)\in\Theta\}$$

Then Θ is clearly the sum of sets H_1, H_2 and B_{12}. Next let $E^{(i)}$ be i-th experiment, and define x_i as follows:

Received April 23, 1964.

$$x_i=1 \qquad \text{if } S \text{ occurs under } E^{(i)},$$
$$=0 \qquad \text{if } F \text{ occurs under } E^{(i)}.$$

In the following line we shall assume that in $E^{(i)}$ S or F occurs independently of the selection of $E^{(1)}, \cdots, E^{(i)}$ $(i=1, 2, \cdots)$. Then we see that $x_1, x_2, \cdots, x_n, \cdots$ are independent random variables. And let n_1 be the number of selections of experiment E_1 in the partial n experiments $E^{(1)}, \cdots, E^{(n)}$, m_1 the number of occurrences of S in these n_1 observations by E_1, and similarly n_2 the number of selections of E_2 in the partial n experiments, and m_2 the number of occurrences of S in these n_2 observations by E_2. Then if $\theta=(p_1, p_2)$ is an element of Θ, the probability density function of x_i at $E^{(i)}$ $f(x_i, \theta, E^{(i)})$ is known to be following form:

$$f(x_i, \theta, E^{(i)})=p_1^{x_i}(1-p_1)^{1-x_i} \qquad \text{if} \quad E^{(i)}=E_1,$$
$$=p_2^{x_i}(1-p_2)^{1-x_i} \qquad \text{if} \quad E^{(i)}=E_2.$$

Then the likelihood function of θ over the partial n experiments is given by $\prod_{i=1}^{n} f(x_i, \theta, E^{(i)})$. This is a function of n observations $x_1, \cdots, x_n$, n experiments $E^{(1)}, \cdots, E^{(n)}$ and θ. The maximum likelihood estimate $\hat{\theta}_n$ of θ over the partial n experiments is not only a function of n observations $x_1, \cdots, x_n$ but also a function of n experiments $E^{(1)}, \cdots, E^{(n)}$. Next we shall denote by $\tilde{\theta}_n$ the maximum likelihood estimate of θ on the closed domain $a(\hat{\theta}_n)$ over the n experiments $E^{(1)}, \cdots, E^{(n)}$, where $a(\hat{\theta}_n)$ is defined as follows:

$$\text{if} \quad \hat{\theta}_n \in H_i, \quad \text{then} \quad a(\hat{\theta}_n)=\Theta-H_i \qquad (i=1, 2)$$
$$\text{and if} \quad \hat{\theta}_n \in B_{12}, \quad \text{then} \quad a(\hat{\theta}_n)=\Theta.$$

Definition of discrimination. As a measure of discrimination between two probability density functions f_1 and f_2, Kullback [3] introduced following

$$I(f_1, f_2)=\int f_1 \log \frac{f_1}{f_2}\, d\mu.$$

In our case, we can use this measure to express the discrimination between $f(x, \theta, E)$ and $f(x, \varphi, E)$, i.e;

$$I((p_1, p_2), (p_1^*, p_2^*), E_1)=p_1 \log \frac{p_1}{p_1^*}+(1-p_1)\log \frac{1-p_1}{1-p_1^*},$$

and

$$I((p_1, p_2), (p_1^*, p_2^*), E_2)=p_2 \log \frac{p_2}{p_2^*}+(1-p_2)\log \frac{1-p_2}{1-p_2^*},$$

where $\theta=(p_1, p_2)$ and $\varphi=(p_1^*, p_2^*)$.

Definition of procedure $\mathfrak{P}$. We shall call procedure $\mathfrak{P}$, if the following conditions are satisfied: $E^{(1)}=E_1$, $E^{(2)}=E_2$ and for $n\geqq 2$ succesively

$$E^{(n+1)} = E_1 \quad \text{if} \quad \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_1)}{c_1} > \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{c_2},$$

(2. 1)
$$= E_2 \quad \text{if} \quad \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_1)}{c_1} < \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{c_2},$$

$$= E^{(n)} \quad \text{if} \quad \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_1)}{c_1} = \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{c_2}.$$

3. Theorems and the proofs.

At first, put

$$D(\theta) = \{(p_1^*, p_2^*): (p_1^* \geqq p_1 \text{ and } p_2^* \leqq p_2) \text{ or } (p_1^* \leqq p_1 \text{ and } p_2^* \geqq p_2)\}$$

where $\theta = (p_1, p_2)$,

(3. 1)
$$\theta_n^* = \left\{\varphi: \frac{I(\hat{\theta}_n, \varphi, E_1)}{c_1} = \frac{I(\hat{\theta}_n, \varphi, E_2)}{c_2}, \ \varphi \in D(\hat{\theta}_n)\right\} \cap B_{12},$$

$$\theta^* = \left\{\varphi: \frac{I(\theta, \varphi, E_1)}{c_1} = \frac{I(\theta, \varphi, E_2)}{c_2}, \ \varphi \in D(\theta)\right\} \cap B_{12} \text{ [1)]}$$

and $\theta_n^* = (p_n^*, p_n^*)$, $\theta^* = (p^*, p^*)$. Using these p_n^*, p^*, we define

(3. 2)
$$\lambda_n^* = \frac{p_n^* - \dfrac{m_2}{n_2}}{\dfrac{m_1}{n_1} - \dfrac{m_2}{n_2}}$$

and

(3. 3)
$$\lambda^* = \frac{p^* - p_2}{p_1 - p_2}.$$

Moreover for fixed λ $\lambda \in [0, 1]$, let $\tilde{\theta}$ be $\tilde{\theta} = (p, p)$, where

(3. 4)
$$p = \lambda(p_1 - p_2) + p_2.$$

Then we can list the following Theorems.

THEOREM 1. *Our procedure* $\mathfrak{L}$ *satisfies the next relation*:

$$\lim_{n \to \infty} \frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}} = I^*(\theta)$$

with probability 1, *where*

1) It is clear that θ_n^*, θ^* are uniquely determined.

$$I^*(\theta)=\frac{I(\theta, \theta^*, E_1)}{c_1}=\frac{I(\theta, \theta^*, E_2)}{c_2} \qquad (\theta=(p_1, p_2)),$$

$$S_n(\hat{\theta}_n, \tilde{\theta}_n)=\log\frac{\prod_{i=1}^n f(x_i, \hat{\theta}_n, E^{(i)})}{\prod_{i=1}^n f(x_i, \tilde{\theta}_n, E^{(i)})}$$

and $C^{(i)}$ *is the cost of* $E^{(i)}$.

THEOREM 2. *Any sequence of experiments* $E^{(n)}$ $(n=1, 2, \cdots)$ *such that* $\lim_{n\to\infty}(n_1/n)=\lambda^*$ *satisfies also the same result as Theorem* 1, *that is*;

$$\lim_{n\to\infty}\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=I^*(\theta)$$

with probability 1.

THEOREM 3. *Given any sequence of experiments* $E^{(n)}$ $(n=1, 2, \cdots)$ *such that* $\lim_{n\to\infty}(n_1/n)=\lambda$ $(\lambda\in[0, 1])$ *and if* $\lim_{n\to\infty}\min(n_1, n_2)=+\infty$, *the next limit*

$$\lim_{n\to\infty}\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=\frac{\lambda I(\theta, \tilde{\theta}, E_1)+(1-\lambda)I(\theta, \tilde{\theta}, E_2)}{\lambda c_1+(1-\lambda)c_2} \tag{3. 5}$$

exists with probability 1.

THEORFM 4. *The limit function of* λ (3. 5) $(\lambda\in[0, 1])$ *has only one maximum value if and only if* $\lambda=\lambda^*$.

In order to prove these theorems we need the following Lemmas.

LEMMA 1. *If we execute any procedure, we have always*

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

and

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

Proof. As the function $p_1^{m_1}(1-p_1)^{n_1-m_1}$ has the maximum value at $p_1=m_1/n_1$ and $p_2^{m_2}(1-p_2)^{n_2-m_2}$ at m_2/n_2, the likelihood function

$$p_1^{m_1}(1-p_1)^{n_1-m_1}p_2^{m_2}(1-p_2)^{n_2-m_2}$$

has the maximum value on Θ at $\hat{\theta}_n=(m_1/n_1, m_2/n_2)$. This $\hat{\theta}_n$ is the maximum likelihood estimate over Θ of $p_1^{m_1}(1-p_1)^{n_1-m_1}p_2^{m_2}(1-p_2)^{n_1-m_2}$.

Next we suppose that

$$\hat{\theta}_n\in H_1 \qquad \text{and} \qquad \tilde{\theta}_n\in H_2.$$

Then as the line $\overline{\hat{\theta}_n\tilde{\theta}_n}$ connecting $\hat{\theta}_n$ and $\tilde{\theta}_n$ crosses B_{12}, we have a crossing point

θ different from $\tilde{\theta}_n \in H_2$. Since $p_1^{m_1}(1-p_1)^{n_1-m_1}$ is monotonically increasing in $(0, m_1/n_1)$ and monotonically decreasing in $(m_1/n_1, 1)$, and also $p_2^{m_2}(1-p_2)^{n_2-m_2}$ is monotonically increasing in $(0, m_2/n_2)$ and monotonically decreasing in $(m_2/n_2, 1)$, it is clear that

$$\prod_{i=1}^{n} f(x_i, \theta, E^{(i)}) > \prod_{i=1}^{n} f(x_i, \tilde{\theta}_n, E^{(i)}).$$

As $\theta \in a(\hat{\theta}_n) = H_2 \cup B_{12}$, the above inequality is contradiction to the definition of $\tilde{\theta}_n$. Thus we can conclude that if $\hat{\theta}_n \in H_1$, then $\tilde{\theta}_n \in B_{12}$.

In the same manner we can show that if $\hat{\theta}_n \in H_2$, then $\tilde{\theta}_n \in B_{12}$ and if $\hat{\theta}_n \in B_{12}$ then $\tilde{\theta}_n \in B_{12}$ and $\hat{\theta}_n = \tilde{\theta}_n$. Hence we see $\tilde{\theta}_n \in B_{12}$ for all cases. Therefore, to find $\tilde{\theta}_n$, we search only on B_{12} so that the likelihood function on B_{12} becomes

$$p^{m_1}(1-p)^{n_1-m_1} p^{m_2}(1-p)^{n_2-m_2} = p^{m_1+m_2}(1-p)^{n_1+n_2-(m_1+m_2)}.$$

Then the function has only one maximum value if and only if

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

Hence

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

LEMMA 2. *Given the sequence of experiments under the procedure* $\mathfrak{L}$ $E^{(1)}, E^{(2)}, \cdots, E^{(n)}, \cdots$ *Then the probability that*

$$E^{(n)} = E_1$$

for all $n \geqq k$ *or*

$$E^{(n)} = E_2$$

for all $n \geqq k$ *is zero, where* k *is any fixed positive integer.*

Proof. Suppose the probability

$$E^{(n)} = E_1$$

for all $n \geqq k$ is positive, where k is any fixed positive integer. Then we have

$$E^{(n)} = E_1$$

and

$$\frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_1)}{c_1} \geqq \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{c_2} \tag{3.6}$$

for all n $(n \geqq k)$, with positive probability. Hence by the law of large numbers, we have

$$\lim_{n\to\infty} \frac{m_1}{n_1} = p_1$$

with positive probability. On the other hand, n_2 and m_2 are invariant for all experiments $E^{(n)}$ $(n\geqq k)$ with positive probability. Hence m_2/n_2 is fixed at the value $[m_2/n_2]_{n=k}$ which is determined by $E^{(1)}, E^{(2)}, \cdots, E^{(k)}$. Hence we see

$$\lim_{n\to\infty}\frac{m_2}{n_2}=\left[\frac{m_2}{n_2}\right]_{n=k}$$

with positive probability. Therefore we have

$$\lim_{n\to\infty}\hat{\theta}_n=\lim_{n\to\infty}\left(\frac{m_1}{n_1},\left[\frac{m_2}{n_2}\right]_{n=k}\right)=\left(p_1,\left[\frac{m_2}{n_2}\right]_{n=k}\right)$$

and

$$\lim_{n\to\infty}\tilde{\theta}_n=\lim_{n\to\infty}\left(\frac{m_1+m_2}{n_1+n_2},\frac{m_1+m_2}{n_1+n_2}\right)=(p_1,p_1)$$

with positive probability. Using these facts, we have

$$\lim_{n\to\infty}\frac{I(\hat{\theta}_n,\tilde{\theta}_n,E_1)}{c_1}=\frac{p_1\log\dfrac{p_1}{p_1}+(1-p_1)\log\dfrac{1-p_1}{1-p_1}}{c_1}=0$$

and

$$\lim_{n\to\infty}\frac{I(\hat{\theta}_n,\tilde{\theta}_n,E_2)}{c_2}=\frac{\left[\dfrac{m_2}{n_2}\right]\log\dfrac{\left[\dfrac{m_2}{n_2}\right]}{p_1}+\left(1-\left[\dfrac{m_2}{n_2}\right]\right)\log\dfrac{1-\left[\dfrac{m_2}{n_2}\right]}{1-p_1}}{c_2}\geqq 0$$

with positive probability, where $[m_2/n_2]$ means $[m_2/n_2]_{n=k}$. Therefore

$$0=\lim_{n\to\infty}\frac{I(\hat{\theta}_n,\tilde{\theta}_n,E_1)}{c_1}\leqq\lim_{n\to\infty}\frac{I(\hat{\theta}_n,\tilde{\theta}_n,E_2)}{c_2}$$

with positive probability. Hence, by (3. 6),

$$0=\lim_{n\to\infty}\frac{I(\hat{\theta}_n,\tilde{\theta}_n,E_1)}{c_1}=\lim_{n\to\infty}\frac{I(\hat{\theta}_n,\tilde{\theta}_n,E_2)}{c_2}$$

with positive probability. It follows clearly that $p_1=[m_2/n_2]_{n=k}$ and hence $m_2/n_2=p_1$ for all n $(n\geqq k)$ with positive probability. Hence, by (3. 6)

$$\frac{1}{c_1}\left\{\frac{m_1}{n_1}\log\frac{\dfrac{m_1}{n_1}}{\dfrac{m_1+m_2}{n_1+n_2}}+\left(1-\frac{m_1}{n_1}\right)\log\frac{1-\dfrac{m_1}{n_1}}{1-\dfrac{m_1+m_2}{n_1+n_2}}\right\}$$

(3. 7)

$$\geqq \frac{1}{c_2}\left\{p_1 \log \frac{p_1}{\frac{m_1+m_2}{n_1+n_2}} +(1-p_1) \log \frac{1-p_1}{1-\frac{m_1+m_2}{n_1+n_2}}\right\}$$

for all n $(n \geqq k)$ with positive probability. Here we consider two functions:

$$f(x)=\frac{1}{c_1}\left\{x \log \frac{x}{\frac{m_1+m_2}{n_1+n_2}} +(1-x) \log \frac{1-x}{1-\frac{m_1+m_2}{n_1+n_2}}\right\}$$

and

$$g(y)=\frac{1}{c_2}\left\{y \log \frac{y}{\frac{m_1+m_2}{n_1+n_2}} +(1-y) \log \frac{1-y}{1-\frac{m_1+m_2}{n_1+n_2}}\right\}$$

and we use Taylor's expansion for $f(x)$ and $g(y)$ around $(m_1+m_2)/(n_1+n_2)$ as follows.

$$f(x)=\frac{1}{2}\left(x-\frac{m_1+m_2}{n_1+n_2}\right)^2\left(\frac{1}{\xi_1}+\frac{1}{1-\xi_1}\right),$$

$$g(y)=\frac{1}{2}\left(y-\frac{m_1+m_2}{n_1+n_2}\right)^2\left(\frac{1}{\xi_2}+\frac{1}{1-\xi_2}\right),$$

where ξ_1: $x<\xi_1<\frac{m_1+m_2}{n_1+n_2}$ or $\frac{m_1+m_2}{n_1+n_2}<\xi_1<x$,

and ξ_2: $y<\xi_2<\frac{m_1+m_2}{n_1+n_2}$ or $\frac{m_1+m_2}{n_1+n_2}<\xi_2<y$.

Then the inequality (3. 7) becomc as follows:

$$\frac{\left(\frac{m_1}{n_1}-\frac{m_1+m_2}{n_1+n_2}\right)^2\left(\frac{1}{\xi_1}+\frac{1}{1-\xi_1}\right)}{\left(p_1-\frac{m_1+m_2}{n_1+n_2}\right)^2\left(\frac{1}{\xi_2}+\frac{1}{1-\xi_2}\right)} \geqq \frac{c_1}{c_2}>0 \tag{3. 8}$$

for all n $(n \geqq k)$, with positive probability. And if $n \to \infty$ then $\xi_1 \to p_1$, $\xi_2 \to p_1$ and

$$\frac{\left(\frac{m_1}{n_1}-\frac{m_1+m_2}{n_1+n_2}\right)^2}{\left(p_1-\frac{m_1+m_2}{n_1+n_2}\right)^2}=\frac{\left(\frac{m_1}{n_1}-p_1\right)^2\left(\frac{n_2}{n}\right)^2}{\left(\frac{m_1}{n_1}-p_1\right)^2\left(\frac{n_1}{n}\right)^2}=\left(\frac{n_2}{n_1}\right)^2 \to 0$$

with positive probability.

Hence, this contradicts (3. 8). Thus we proved the probability that

$$E^{(n)}=E_1$$

exists for all n $(n \geqq k)$ is zero, where k is any fixed positive integer.

In the same manner, we can provφ that the probability

$$E^{(n)}=E_2$$

for all n ($n \geqq m$) is zero, where m is any fixed positive integer.

LEMMA 3. *Given the sequence of experiments under the procedure* $\mathfrak{P}$

$$E^{(1)}, E^{(2)}, \cdots, E^{(n)}, \cdots,$$

then we have

$$P\{\min(n_1, n_2) \to \infty \quad \text{as } n \to \infty\} = 1.$$

Proof. Suppose that exists a constant k such that as $n \to \infty$

$$\min(n_1, n_2) \leqq k$$

with positive probability. Then we have

$$E^{(n)}=E_1$$

for all n ($n \geqq m$) or

$$E^{(n)}=E_2$$

for all n ($n \geqq m$) with positive probability, where m is a fixed positive integer. But by Lemma 2 we know that these facts do not exist.

LEMMA 4. *We have*

$$\lim_{n\to\infty} \hat{\theta}_n = (p_1, p_2)$$

with probability 1.

Proof, By Lemma 3, we know that if $n \to \infty$ then $n_1 \to \infty$ and $n_2 \to \infty$ with probability 1. Hence, using the law of large numbers,

$$\lim_{n\to\infty} \frac{m_1}{n_1} = p_1 \text{ and } \lim_{n\to\infty} \frac{m_2}{n_2} = p_2, \text{[2)]}$$

with probability 1. Therefore

$$\lim_{n\to\infty} \hat{\theta}_n = (p_1, p_2)$$

with probability 1.

LEMMA 5. *We have*

$$\lim_{n\to\infty} \frac{n_1}{n} = \lambda^*$$

2) See, for example, Halmos [2].

with probability 1, *where* λ^* *is defined by* (3. 3).

Proof. Evidently, we have

$$\lim_{n\to\infty} \theta_n{}^* = \theta^*$$

with probability 1, by Lemma 4. Hence by the definition of $\lambda_n{}^*$

$$\lim_{n\to\infty} \lambda_n{}^* = \lambda^* \tag{3. 9}$$

with probability 1. It is easily verified that the procedure $\mathfrak{L}$ is equivalent to the following conditions:

$$E^{(1)} = E_1, \qquad E^{(2)} = E_2$$

and for $n \geqq 2$

$$\begin{aligned} E^{(n+1)} &= E_1 && \text{if } \frac{n_1}{n} < \lambda_n{}^*, \\ &= E_2 && \text{if } \frac{n_1}{n} > \lambda_n{}^*, \\ &= E^{(n)} && \text{if } \frac{n_1}{n} = \lambda_n{}^* \end{aligned} \tag{3. 10}$$

respectively. Using this property of the procedure $\mathfrak{L}$, we shall show

$$\lim_{n\to\infty} \frac{n_1}{n} = \lambda^*$$

with probability 1. For any positive number ε, by (3. 9), there exists some integer n_0 such that

$$|\lambda_n{}^* - \lambda^*| < \frac{\varepsilon}{2}$$

for all n ($n \geqq n_0$) with probability 1.

Now we consider the following two cases:

$$\text{(i)} \quad \left| \left[\frac{n_1}{n} \right]_{n=n_0} - \lambda^* \right| \geqq \frac{\varepsilon}{2}$$

with probability 1 and

$$\text{(ii)} \quad \left| \left[\frac{n_1}{n} \right]_{n=n_0} = \lambda^* \right| < \frac{\varepsilon}{2}$$

with probability 1, where $[n_1/n]_{n=n_0}$ is the relative frequency of selection of E_1 from $E^{(1)}$ to $E^{(n_0)}$.

If (i), by the property (3. 10) of $\mathfrak{L}$, there exists j_0 ($j_0 \geqq n_0$)

$$\left|\left[\frac{n_1}{n}\right]_{n=j_0} - \lambda^*\right| < \frac{\varepsilon}{2} \tag{3. 11}$$

with probability 1.

If (ii), we put $j_0 = n_0$ and we have (3. 11) with probability 1. Hence, we can find the first integer j_0 ($j_0 \geqq n_0$) satisfying (3. 11). Next we suppose that there exists an integer k ($k \geqq j_0$) such that

$$\left|\left[\frac{n_1}{n}\right]_{n=k-1} - \lambda^*\right| < \frac{\varepsilon}{2}$$

and

$$\left|\left[\frac{n_1}{n}\right]_{n=k} - \lambda^*\right| \geqq \frac{\varepsilon}{2} \tag{3. 12}$$

with probability 1. Then by the procedure $\mathfrak{L}$, we see

$$\left|\left[\frac{n_1}{n}\right]_{n=k} - \lambda^*\right| > \left|\left[\frac{n_1}{n}\right]_{n=k+1} - \lambda^*\right|$$

with probability 1. Since generally, the fact

$$\left|\left[\frac{n_1}{n}\right]_{n=k} - \left[\frac{n_1}{n}\right]_{n=k-1}\right| < \frac{2}{k}$$

with probability 1 is satisfied for all k, we see

$$\left|\left[\frac{n_1}{n}\right]_{n=k} - \lambda^*\right|$$

$$\leqq \left|\left[\frac{n_1}{n}\right]_{n=k} - \left[\frac{n_1}{n}\right]_{n=k-1}\right| + \left|\left[\frac{n_1}{n}\right]_{n=k-1} - \lambda^*\right| < \frac{2}{k} + \frac{\varepsilon}{2}$$

with probability 1. Therefore if we have (3. 12) with probability 1, we have

$$\left|\left[\frac{n_1}{n}\right]_{n=k} - \lambda^*\right| < \frac{2}{k} + \frac{\varepsilon}{2} \leqq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

with probability 1 for any k ($k \geqq j_0$) such as

$$\frac{2}{k} \leqq \frac{\varepsilon}{2},$$

and if there does not exist k ($k \geqq j_0$) satisfying (3. 12) with probability 1, we have

$$\left|\left[\frac{n_1}{n}\right]_{n=k} - \lambda^*\right| < \frac{\varepsilon}{2} < \varepsilon$$

with probability 1 for all k ($k \geqq j_0$). Thus we have

$$\left|\frac{n_1}{n} - \lambda^*\right| < \varepsilon$$

with probability 1 for all n ($n \geqq N_0$), where

$$N_0 = \max\left[\frac{4}{\varepsilon}, j_0\right].$$

Thus, $\lim_{n\to\infty}(n_1/n) = \lambda^*$ with probability 1, as to be proved.

LEMMA 6. $\lim_{n\to\infty}\tilde{\theta}_n = \theta^*$ *with prabability* 1.

Proof. By simple calculation, we have

$$\frac{m_1+m_2}{n_1+n_2} = \frac{n_1}{n}\left(\frac{m_1}{n_1} - \frac{m_2}{n_2}\right) + \frac{m_2}{n_2}$$

and, from Lemma 4,

$$\lim_{n\to\infty}\frac{m_1}{n_1} = p_1$$

and

$$\lim_{n\to\infty}\frac{m_2}{n_2} = p_2$$

with probability 1, and, by Lemma 5, we see that

$$\lim_{n\to\infty}\frac{n_1}{n} = \lambda^*$$

with probability 1. Then

$$\lim_{n\to\infty}\frac{m_1+m_2}{n_1+n_2} = \lim_{n\to\infty}\left\{\frac{n_1}{n}\left(\frac{m_1}{n_1} - \frac{m_2}{n_2}\right) + \frac{m_2}{n_2}\right\}$$

$$= \lambda^*(p_1 - p_2) + p_2 = p^*$$

with probability 1 by the definition of λ^*. Hence,

$$\lim_{n\to\infty}\tilde{\theta}_n = \theta^* = (p^*, p^*)$$

with probability 1, as to be proved.

Proof of Theorem 1.

$$S_n(\hat{\theta}_n, \tilde{\theta}_n) = \log\frac{\prod_{i=1}^n f(x_i, \hat{\theta}_n, E^{(i)})}{\prod_{i=1}^n f(x_i, \tilde{\theta}_n, E^{(i)})}$$

$$=\log\frac{\left(\frac{m_1}{n_1}\right)^{m_1}\left(1-\frac{m_1}{n_1}\right)^{n_1-m_1}\left(\frac{m_2}{n_2}\right)^{m_2}\left(1-\frac{m_1}{n_1}\right)^{n_2-m_2}}{\left(\frac{m_1+m_2}{n}\right)^{m_1}\left(1-\frac{m_1+m_2}{n}\right)^{n_1-m_1}\left(\frac{m_1+m_2}{n}\right)^{m_2}\left(1-\frac{m_1+m_2}{n}\right)^{n_2-m_2}}$$

$$=m_1\log\frac{\frac{m_1}{n_1}}{\frac{m_1+m_2}{n}}+(n_1-m_1)\log\frac{1-\frac{m_1}{n_1}}{1-\frac{m_1+m_2}{n}}+m_2\log\frac{\frac{m_2}{n_2}}{\frac{m_1+m_2}{n}}$$

$$+(n_2-m_2)\log\frac{1-\frac{m_2}{n_2}}{1-\frac{m_1+m_2}{n}}$$

$$=n_1\left\{\frac{m_1}{n_1}\log\frac{\frac{m_1}{n_1}}{\frac{m_1+m_2}{n}}+\left(1-\frac{m_1}{n_1}\right)\log\frac{1-\frac{m_1}{n_1}}{1-\frac{m_1+m_2}{n}}\right\}$$

$$+n_2\left\{\frac{m_2}{n_2}\log\frac{\frac{m_2}{n_2}}{\frac{m_1+m_2}{n}}+\left(1-\frac{m_2}{n_2}\right)\log\frac{1-\frac{m_2}{n_2}}{1-\frac{m_1+m_2}{n}}\right\}$$

$$=n_1 I(\hat{\theta}_n, \tilde{\theta}_n, E_1)+n_2 I(\hat{\theta}_n, \tilde{\theta}_n, E_2).$$

Hence

$$\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=\frac{n_1c_1}{n_1c_1+n_2c_2}\,\frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_1)}{c_1}+\frac{n_2c_2}{n_1c_1+n_2c_2}\,\frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{c_2}$$

Therefore, as $\lim_{n\to\infty}\hat{\theta}_n=\theta$, $\lim_{n\to\infty}\tilde{\theta}_n=\theta^*$ with probability 1, from the Lemma 4 and the Lemma 6, we have

$$\lim_{n\to\infty}\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=I^*(\theta)$$

with probability 1.

Proof of Theorem 2. Any sequence of experiments $E^{(n)}$ $(n=1, 2, \cdots)$ such that $\lim_{n\to\infty}n_1/n=\lambda^*$, satisfy the Lemma 4 and Lemma 6 evidently. Hence

$$\lim_{n\to\infty}\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=I^*(\theta)$$

with probability 1.

Proof of Theorem 3. It is clear by the hypothesis

$$\lim_{n\to\infty} \min (n_1, n_2) = +\infty$$

that $\lim_{n\to\infty}\hat{\theta}_n=\theta$ with probability 1. And using the equality.

$$\frac{m_1+m_2}{n}=\frac{n_1}{n}\left(\frac{m_1}{n_1}-\frac{m_2}{n_2}\right)+\frac{m_2}{n_2}$$

we have

$$\lim_{n\to\infty}\frac{m_1+m_2}{n}=\lambda(p_1-p_2)+p_2=p$$

with probability 1. Then, we have

$$\lim_{n\to\infty}\tilde{\theta}_n=\tilde{\theta}=(p, p)$$

with probability 1. Therefore, we see easily that

$$\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=\frac{n_1 I(\hat{\theta}_n, \tilde{\theta}_n, E_1)+n_2 I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{n_1 c_1+n_2 c_2}$$

$$=\frac{\frac{n_1}{n} I(\hat{\theta}_n, \tilde{\theta}_n, E_1)+\left(1-\frac{n_1}{n}\right) I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{\frac{n_1}{n} c_1+\left(1-\frac{n_1}{n}\right)c_2}$$

Hence, we have

$$\lim_{n\to\infty}\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=\frac{\lambda I(\theta, \tilde{\theta}, E_1)+(1-\lambda)I(\theta, \tilde{\theta}, E_2)}{\lambda c_1+(1-\lambda)c_2}$$

with probability 1, as to be proved.

Proof of Theorem 4. As p was defined as $\lambda(p_1-p_2)+p_2$ in (3. 4), we have

$$\lambda=\frac{p-p_2}{p_1-p_2}.$$

Hence, by simple calculation, we have

$$\frac{d}{dp}\left\{\frac{\lambda I(\theta, \tilde{\theta}, E_1)+(1-\lambda)I(\theta, \tilde{\theta}, E_2)}{\lambda c_1+(1-\lambda)c_2}\right\}$$

$$=\frac{c_1\cdot c_2}{\{\lambda c_1+(1-\lambda)c_2\}^2}\;\frac{1}{p_1-p_2}\left\{\frac{I(\theta, \tilde{\theta}, E_1)}{c_1}-\frac{I(\theta, \tilde{\theta}, E_2)}{c_2}\right\}$$

Therefore the derivative is equal to zero if and only if $\tilde{\theta}=\theta^*$. Thus, the function of λ

$$\frac{\lambda I(\theta, \tilde{\theta}, E_1)+(1-\lambda)I(\theta, \tilde{\theta}, E_2)}{\lambda c_1+(1-\lambda)c_2}$$

has only one maximum value if and only if $\lambda=\lambda^*$, because $\tilde{\theta}=\theta^*$ is equivalent to $\lambda=\lambda^*$.

ACKNOWLEDGEMENT. The author expresses hearty thanks to Professor K. Kunisawa for useful suggestions.

REFERENCES

[1] CHERNOFF, H., Sequential design of experiments. Ann. Math. Stat. **30** (1959), 755-770.
[2] HALMOS, P. R., Measure theory. (1959), Van Nostrand.
[3] KULLBACK, S., Information theory and statistics. (1959), Wiley.
[4] KUNISAWA, K., Introduction to information theory for operations research. 4th Ed. (1963), J.U.S.E. (In Japanese)

DEPARTMENT OF MATHEMATICS,
TOKYO INSTITUTE OF TECHNOLOGY.