ASYMPTOTIC BEHAVIOR OF SEQUENTIAL DESIGN WITH COSTS OF EXPERIMENTS

(THE CASE OF NORMAL DISTRIBUTION)

BY KAZUTOMO KAWAMURA

1. Introduction.

Recently we showed the following fact in [2]. For two binomial trials E_i $(i=1, 2)$ with parameter p_i and cost c_i a procedure $\mathfrak{P}$ is given such that, using the procedure, the sum of information discriminating two hypotheses $p_1>p_2$ and $p_1<p_2$ per unit cost is asymptotically maximized.

In this paper we shall consider two kinds of normal trials E_i $(i=1, 2)$ which have mean m_i, same variance σ^2 and cost c_i where we assume $m_1 \neq m_2$. The object of this paper is also to discriminate the hypotheses $m_1>m_2$ or $m_1<m_2$. Then the same procedure as in [2] is optimal and we shall show analogously the asymptotic behavior of the procedure.

2. Notations and definitions.

We shall denote by Θ 2-dimensional euclidean parameter space. An element of of the space Θ is expressed by $\theta=(m_1, m_2)$ and we put

$$H_1=\{(m_1, m_2): m_1>m_2, (m_1, m_2)\in\Theta\},$$

$$H_2=\{(m_1, m_2): m_1<m_2, (m_1, m_2)\in\Theta\}$$

and

$$B=\{(m_1, m_2): m_1=m_2, (m_1, m_2)\in\Theta\}.$$

Then $\Theta=H_1+H_2+B$ is satisfied. Next let $E^{(i)}$ be i-th experiment, and define X_i to be i-th random variable given by $E^{(i)}$. We assume that X_{i+1} occurs in $E^{(i+1)}$ independently of the selection of $E^{(1)}, E^{(2)}, \cdots, E^{(i)}$ $(i=1, 2, \cdots)$. 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 $E^{(1)}, E^{(2)}, \cdots, E^{(n)}$ and similarly n_2 the number of selections of E_2 in the partial n experiments. If $\theta=(m_1, m_2)$ 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:

Received September 10, 1964.

$$f(x_i, \theta, E^{(i)})=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x_i-m_1)^2}{2\sigma^2}\right\} \qquad \text{if} \quad E^{(i)}=E_1,$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x_i-m_2)^2}{2\sigma^2}\right\} \qquad \text{if} \quad E^{(i)}=E_2.$$

Then the likelihood function of θ over the partial n experiments is given by $\Pi_{i=1}^n f(x_i, \theta, E^{(i)})$. This is a function of n observations $x_1, x_2, \cdots, x_n$, n experiments $E^{(1)}, E^{(2)}, \cdots, E^{(n)}$ and θ. The maximum likelihood estimate $\hat{\theta}_n$ of θ over the partial n experiments is not only a function of n observations $x_1, x_2, \cdots, x_n$, but also a function of n experiments $E^{(1)}, E^{(2)}, \cdots, E^{(n)}$. Next we shall denote by $\tilde{\theta}_n$ the maximum likelihood estimate of θ on the subspace $a(\hat{\theta}_n)$ over n experiments $E^{(1)}, E^{(2)}, \cdots, E^{(n)}$ where $a(\hat{\theta}_n)$ is defined as follows:

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

$$=\Theta \qquad \text{if} \quad \hat{\theta}_n \in B.$$

Definition of discrimination. As a measure of discrimination between two probability density functions f_1, f_2 Kullback [4] introduced

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

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

$$I(\theta, \varphi, E_1)=\int_R\left[\log\frac{f(x, \theta, E_1)}{f(x, \varphi, E_1)}\right]f(x, \theta, E_1)\, dx=\frac{(m_1-m_1^*)^2}{2\sigma^2}. \tag{2. 1}$$

Similarly we can verify

$$I(\theta, \varphi, E_2)=\frac{(m_2-m_2^*)^2}{2\sigma^2} \tag{2. 2}$$

where $\theta=(m_1, m_2)$ and $\varphi=(m_1^*, m_2^*)$.

Definition of procedure $\mathfrak{L}$. We shall call next policy procedure $\mathfrak{L}$: $E^{(1)}=E_1$, $E^{(2)}=E_2$ and for $n\geqq 2$ succesively

$$E^{(n+1)}=\begin{Bmatrix} E_1 \\ E_2 \\ E^{(n)} \end{Bmatrix} \quad \text{if} \quad \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_1)}{c_1}\begin{Bmatrix} > \\ < \\ = \end{Bmatrix}\frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{c_2}. \tag{2. 3}$$

3. Theorems.

First we put

$$D(\theta)=\{(m_1^*, m_2^*): (m_1^* \geqq m_1, m_2^* \leqq m_2) \text{ or } (m_1^* \leqq m_1, m_2^* \geqq m_2)\},$$

where $\theta=(m_1, m_2)$, and

$$\theta^*=\left\{\varphi: \frac{I(\theta, \varphi, E_1)}{c_1}=\frac{I(\theta, \varphi, E_2)}{c_2}, \ \varphi \in D(\theta)\right\} \cap B=(m^*, m^*). \tag{3.1}$$

Using this m^*, we define

$$\lambda^*=\frac{m^*-m_2}{m_1-m_2}. \tag{3.2}$$

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

$$m=\lambda m_1+(1-\lambda)m_2. \tag{3.3}$$

Then we can list the following Theorems.

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

$$\lim_{n\to\infty} \frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=I^*(\theta) \tag{3.4}$$

with probability 1, *where*

$$I^*(\theta)=\frac{I(\theta, \theta^*, E_1)}{c_1}=\frac{I(\theta, \theta^*, E_2)}{c_2} \qquad (\theta=(m_1, m_2)), \tag{3.5}$$

$$S_n(\hat{\theta}_n, \tilde{\theta}_n)=\log \frac{\prod_{i=1}^n f(x_i, \hat{\theta}_n, E^{(i)})}{\prod_{i=1}^n f(x_i, \tilde{\theta}_n, E^{(i)})} \tag{3.6}$$

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

$$\lim_{n\to\infty} \frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}} = \frac{\lambda I(\theta, \tilde{\theta}, E_1)+(1-\lambda)I(\theta, \tilde{\theta}, E_2)}{\lambda c_1+(1-\lambda)c_2} \tag{3.7}$$

exists with probability 1.

THEOREM 4. *The limit function* (3. 7) *of* $\lambda\in[0, 1]$ *has only one maximum value if and only if* $\lambda=\lambda^*$.

Theorems 1~4 explain that the procedure $\mathfrak{L}$ has the property which is asymptotically most informative per unit cost for any other procedure.

In order to prove these Theorems 1~4 we need only the following Lemmas.

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

$$\hat{\theta}_n=\left(\frac{1}{n_1}\sum_{E^{(i)}=E_1} x_i, \frac{1}{n_2}\sum_{E^{(i)}=E_2} x_i\right), \quad \tilde{\theta}_n=\left(\frac{1}{n}\sum_{i=1}^n x_i, \frac{1}{n}\sum_{i=1}^n x_i\right)$$

where $\sum_{E^{(i)}=E_j}$ *means the summation over* $i=1, 2, \cdots, n$ *satisfying* $E^{(i)}=E_j$ $(j=1, 2)$.

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

LEMMA 3. *Under the procedure* $\mathfrak{L}$ *we have*

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

with probability 1.

LEMMA 4. *Under the procedure* $\mathfrak{L}$ *we have*

$$\lim_{n\to\infty} \hat{\theta}_n=(m_1, m_2)$$

with probability 1.

LEMMA 5. *Under the procedure* $\mathfrak{L}$ *we have*

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

with probability 1.

LEMMA 6. *Under the procedure* $\mathfrak{L}$ *we have*

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

with probability 1.

These proofs of Lemma 1∼6 can be lead analogously as given in [2]. But proof of Lemma 5 can be given directly from equivalent condition (3. 8) of procedure $\mathfrak{L}$ as follows.

$$\frac{n_2}{n_1}\begin{Bmatrix} > \\ < \\ = \end{Bmatrix}\sqrt{\frac{c_1}{c_2}}. \tag{3. 8}$$

Then under procedure $\mathfrak{L}$ we can verify that n_1/n approaches to $\sqrt{c_2}/(\sqrt{c_1}+\sqrt{c_2})$ as $n\to\infty$ with probability 1, and from (2. 1), (2. 2), (3. 1) and (3. 2) the limit value $\sqrt{c_2}/(\sqrt{c_1}+\sqrt{c_2})$ of n_1/n is equal to λ^*.

Acknowledgement. The author expresses hearty thanks to Professor K. Kunisawa for many useful suggestions at the seminars in this summer vacation.

REFERENCES

[1] CHERNOFF, H., Sequential design of experiments. Ann. Math. Stat. **30** (1959), 755–770

[2] KAWAMURA, K.,, Asymptotic behavior of sequential design with costs of experiments. Kōdai Math. Sem. Rep. **16** (1964), 169–182.

[3] KAWATA, T., Probability theory and statistics. 5th Ed. (1964), Asakura Co. (In Japanese)

[4] KULLBACK, S., Information theory and statistics. (1959), wiley.

[5] KUNISAWA, K., Modern probability theory. 12th Ed. (1963), Iwanami Co. (In Japanese)

[6] 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.