KŌDAI MATH. SEM. REP.
23 (1971), 208—214

A LIMIT PROPERTY OF SEQUENTIAL DECISION PROCESS

BY KAZUTOMO KAWAMURA

1. Introduction.

In this paper we consider a limit property of sequential design problem. Generally, in the sequential design problem, the following subjects are important for us. (1) Which experiment we must select at each step? (2) When we must stop experiment? Bradt, Johnson and Karlin [1] have given a famous example of sequential design problem, named as "Two Armed Bandit Problem". In the sequential testing model of composite hypothesis, Chernoff [2] has given an asymptotically optimal procedure and its asymptotic behavior.

In this paper we shall give a set of finitely many experiments which are conditioned by the restrictions 1, 2, $\cdots$, 5 in the following section, and give a procedure $\mathcal{P}$ regarding only the subject (1). The process of gains given by the selected experiments being denoted as $\{X_n\}$ under the procedure $\mathcal{P}$, it will be shown that the sample mean $\bar{X}_n=\sum_{i=1}^{n}X_i/n$ given at n-th step has a maximum limit value as n tends to infinity. Under another procedure $\mathcal{P}'$, regarding only the subject (1), the process $\bar{X}_n$ has not always a limit value, but, if the process $\bar{X}_n$ converges under $\mathcal{P}'$, the limit value will be not greater than the limit value under the procedure $\mathcal{P}$.

2. Notations, restrictions of experiments and definition of procedure.

In this paper we treat a set of k mutually independent experiments (trials) $E_1, \cdots, E_k$. The chance variables $X_{E_1}, \cdots, X_{E_k}$ of the trials have unknown mean values $m_1, \cdots, m_k$ respectively. We assume that we have not any information for $m_1, \cdots, m_k$ until we observe the first result of the trials E_1 or E_2 or $\cdots$ or E_k. In the first step of the selection we are admitted to select each one of the given k trials $E_1, \cdots, E_k$. In the second step we are admitted to select the second trial investigating our own purpose under given informations of the result of first step, and so on. And we assume that the result of a fixed step was independently distributed to the preceding selections of trials. Our purpose of selections is to maximize asymptotically the process of the sample mean $\bar{X}_n$ given by the first n observations $X_1, \cdots, X_n$ of the first n selections of trials $E^{(1)}, \cdots, E^{(n)}$. Here we denote the first n observations as $X_1, \cdots, X_n$, and especially we shall not write the trial suffix in the following lines. We shall decide the selecting way of trial for each step

Recived August 31, 1970.

using a posteriori probabilities of Bayes type $\lambda_{n+1,1}, \cdots, \lambda_{n+1,k}$ of hypotheses H_1: $m_1 > \max_{i \neq 1}(m_i), \cdots, H_k$: $m_k > \max_{i \neq k}(m_i)$ given by preceding n observations $X_1, \cdots, X_n$. For $(n+1)$-th step we shall select $(n+1)$-th trial $E^{(n+1)}$ randomizedly from the k trials $E_1, \cdots, E_k$ by selecting ratio $\lambda_{n+1,1}, \cdots, \lambda_{n+1,k}$ of $E_1, \cdots, E_k$ multinomially. Then our process of sample mean $\overline{X}_n$ given by our randomized selections of trials $E^{(1)}, E^{(2)}, \cdots$ has maximum limit value as will be proved in section 3. And we can observe the special case of this model in corollary in section 3. An element E of the given set of trials has a result X which is considered to be a chance variable, and the mean value m of X is unknown value and the density function $f(x, m)$ of X for fixed trial E has following five restrictions.

Restriction 1. We have not any information for the mean m until we observe the result of the trial E.

Restriction 2. The density function $f(x, m)$ is defined on m interval of full line and $f(x, m)$ is m integrable on its domain.

Restriction 3. Our unknown true parameter m is an inner point of its domain interval.

Restriction 4. If we have n observations of independent trials of fixed E denoted as $x_1, \cdots, x_n$ then $\Pi_{i=1}^n f(x_i, m)$ is positive on inner point of m domain and unimodal on the domain.

If we have n results of independent trials of fixed E denoted as $x_1, \cdots, x_n$ then the normalization of the likelihood function $\Pi_{i=1}^n f(x_i, m)$ for our unknown parameter m denoted as $|\Pi_{i=1}^n f(x_i, m)|$ is given as follows:

$$\left|\prod_{i=1}^{n} f(x, m)\right| = \frac{1}{\int \prod_{i=1}^{n} f(x_i, m)dm} \prod_{i=1}^{n} f(x_i, m). \tag{2.1}$$

Restriction 5. The normalization $|\Pi_{i=1}^n f(x_i, m)|$ converges to a m density function which concentrates only at our unknown parameter m with probability one as n tends to infinity.

Before defining our procedure $\mathcal{P}$, we assume that we have $n = n_1 + \cdots + n_k$ observations from n_1 trials of $E_1, \cdots$, from n_k trials of E_k in some way. We shall define $\lambda_{n+1,1}, \cdots, \lambda_{n+1,k}$ after observing n samples $x_1, \cdots, x_n$ as follows:

$$\begin{aligned} \lambda_{n+1,1} &= \int_{H_1} \left|\prod_{i=1}^{n} f(x_i, m^{(i)})\right| dm_1 \cdots dm_k, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \lambda_{n+1,k} &= \int_{H_k} \left|\prod_{i=1}^{n} f(x_i, m^{(i)})\right| dm_1 \cdots dm_k, \end{aligned} \tag{2.2}$$

where $m^{(i)}$ is the mean value corresponding to i-th trial $E^{(i)}$.

Definition of procedure $\mathcal{P}$. We select $(n+1)$-th trial $E^{(n+1)}$ from k trials $E_n, \cdots, E_k$ multinomially with probabilities $\lambda_{n+1,1}, \cdots, \lambda_{n+1,k}$ under the information of n observations $X_1, \cdots, X_n$ until n-th step, which have been defined by (2. 2).

3. Proof of main theorem.

The object of this section is to prove the following main theorem.

THEOREM. *Under our procedure $\mathcal{P}$ the process $(X_1+\cdots+X_n)/n$ converges to the value* $\max(m_1, \cdots, m_k)$ *with probability one as n increases to infinity.*

To prove this theorem it is sufficient to prove the following lemma.

LEMMA. *Under our $\mathcal{P}$, if any trial of the set $E_1, \cdots, E_k$ satisfies the restrictions* 1, $\cdots$, 5, *then the selecting time n_1 of $E_1, \cdots, n_k$ of E_k till n-th step increases to infinity as the number of observations n increases to infinity.*

Proof. To prove that selections of $E_1, \cdots, E_k$ continue infinitly many times as n increases to infinity under our procedure $\mathcal{P}$, contrary to the comment we assume that there exist some trial E_j and an integer N such that $n_j<N$ and $n_i \to \infty$ as $n\to\infty$ $(i\neq j)$. In this contrary assumption, it can be shown that our selecting ratio $\lambda_{n+1,j}$ defined in (2. 2) approaches to

$$\int_{\bar{H}_j}\left|\prod_{E^{(i)}=E_j} f(x_i, m_j)\right| dm_j$$

by the restriction 5, where $\Pi_{E^{(i)}=E_j}$ means the product of n_j functions $f(x_i, m_j)$ satisfying $E^{(i)}=E_j$, $i=1, 2, \cdots, n$, because we can verify

$$\lambda_{n+1,j}=\int_{H_j}\left|\prod_{i=1}^{n} f(x_i, m^{(i)})\right| dm_1 dm_2 \cdots dm_k$$

$$=\int_{H_j}\left|\prod_{E^{(i)}=E_1} f(x_i, m_1)\prod_{E^{(i)}=E_2} f(x_i, m_2)\cdots\prod_{E^{(i)}=E_k} f(x_i, m_k)\right| dm_1 dm_2\cdots dm_k$$

$$\to\int_{H_j}\delta(m_1)\cdots\delta(m_{j-1})\left|\prod_{E^{(i)}=E_j} f(x_i, m_j)\right|\delta(m_{j+1})\cdots\delta(m_k)dm_1dm_2\cdots dm_k$$

$$=\int_{M_j}\left|\prod_{E^{(i)}=E_j} f(x_i, m_j)\right|\left\{\int_{m_j\geqq m_1}\delta(m_1)\,dm_1\cdots\right. \tag{3. 2}$$

$$\left.\int_{m_j\geqq m_{j-1}}\delta(m_{j-1})\,dm_{j-1}\int_{m_j\geqq m_{j+1}}\delta(m_{j+1})dm_{j+1}\cdots\int_{m_j\geqq m_k}\delta(m_k)dm_k\right\}dm_j \quad [1]$$

1) M_j means full range of m_j.

$$=\int_{m_j\geqq\max(\bar{m}_1,\cdots,\bar{m}_{j-1},\bar{m}_{j+1},\cdots,\bar{m}_k)}\left|\prod_{E(i)=E_j}f(x_i,m_j)\right|dm_j$$

$$=\int_{\bar{H}_j}\left|\prod_{E(i)=E_j}f(x_i,m_j)\right|dm_j,$$

where $\delta(m_i)$ is a limit probability density function such that $\int\delta(m_i)dm_i=1$ and $\delta(m_i)=0$ for any value m_i different to unknown true value $\bar{m}_i$ and $\bar{H}_j$ means m_j interval: $m_j\geqq\max(\bar{m}_1,\cdots,\bar{m}_{j-1},\bar{m}_{j+1},\cdots,\bar{m}_k)$.

We can easily see that the limit value (3. 1) is positive by the unimodality given in the restriction 4. Under our procedure $\mathcal{P}$, $\lambda_{n+1,j}$ approaches to the positive value

$$\int_{\bar{H}_j}\left|\prod_{E(i)=E_j}f(x_i,m_j)\right|dm_j \tag{3. 3}$$

as n increases to infinity. Therefore the selecting ratio of E_j approaches to the value. Under our procedure $\mathcal{P}$ the ratio n_j/n approaches to the value, then we have $n_j\to\infty$ as $n\to\infty$. This contradicts to the assumption. Hence under our procedure $\mathcal{P}$ we have $n_j\to+\infty$ for all j as $n\to\infty$. From this fact we get our lemma as to be proved.

Proof of theorem. By the result of the lemma, under our $\mathcal{P}$ we have selecting number n_j of E_j till n-th step converges to $+\infty$ as n increases to infinity, then we have

$$\left|\prod_{E(i)=E_j}f(x_i,m_j)\right|\to\delta(m_j)\quad\text{as}\quad n\to\infty, \tag{3. 4}$$

where $\delta(m_j)$ is a limit probability density function satisfying $\int\delta(m_j)dm_j=1$ and $\delta(m_j)=0$ for all m_j different to unknown true value $\bar{m}_j$. Therefore under our procedure $\mathcal{P}$ we have

$$\lim_{n\to\infty}\left|\prod_{i=1}^{n}f(x,m^{(i)})\right|=\delta(m_1)\delta(m_2)\cdots\delta(m_k) \tag{3. 5}$$

then

$$\begin{aligned}\lambda_{n+1,j}&=\int\cdots\int_{H_j}\left|\prod_{i=1}^{n}f(x_i,m^{(i)})\right|dm_1\cdots dm_k\\ &\to\int\cdots\int_{H_j}\delta(m_1)\cdots\delta(m_k)dm_1\cdots dm_k\end{aligned} \tag{3. 6}$$

the limiting value equals to one if $\bar{m}_j=\max(\bar{m}_1,\cdots,\bar{m}_k)$ and $\bar{m}_j\neq\bar{m}_i$ for all i differing to j and equals to zero if $\bar{m}_j<\max(\bar{m}_1,\cdots,\bar{m}_k)$.

If the set of $\bar{m}_1, \cdots, \bar{m}_k$ has two or more maximum values $\bar{m}_i, \bar{m}_j, \cdots$ then the limit value of the sum of $\lambda_{n+1,i}, \lambda_{n+1,j}, \cdots$, corresponding to $\bar{m}_i, \bar{m}_j, \cdots$, equals to one.

In the following lines, we treat two binomial trials E_1, E_2. That is, we consider E_1, E_2 as binomial trials having following unknown parameters p_1, p_2.

$$P\{X=1|E_i\}=p_i$$
$$P\{X=0|E_i\}=1-p_i, \qquad (i=1, 2). \tag{3.7}$$

For this two trials E_1, E_2, we get following corollary.

COROLLARY. *For this binomial trials E_1, E_2, we additionary assume only following two restrictions:*

Restriction 1. *We have not any information for unknown true parameter $p_\cdot$ until we observe the first observation of trirl $E_\cdot$ ($\cdot=1, 2$ respectively).*

Restriction 3. *Our unknown true parameter point $p_\cdot$ is an inner point of its domain interval* [0, 1].

Then, under our procedure $\mathcal{P}$, we have $\lim_{n\to\infty}(X_1+\cdots+X_n)/n=\max(p_1, p_2)$ *with probability* 1.

Proof of corollary. To prove $\lim_{n\to\infty}(X_1+\cdots+X_n)/n=\max(p_1, p_2)$ with probability 1, we need five restrictions 1, $\cdots$, 5 as proved in our theorem. And in this corollary we assumed two restrictions 1, 3, so that we need to show three restrictions 2, 4 and 5. The definition of binomial trial shows that the domain of probability is expressed by unit interval $0\leqq p\leqq 1$. Terefore we get restriction 2 as to be proved. In the following line we shall show restriction 4. If we have $n_\cdot$ results of $n_\cdot$ independent trials of $E_\cdot$ denoted as $X_1, \cdots, X_{n_\cdot}$. Then

$$\prod_{i=1}^{n_\cdot} P\{x_i|E_\cdot\}=\prod_{i=1}^{n_\cdot} p_\cdot^{x_\cdot}(1-p_\cdot)^{1-x_\cdot} \tag{3.8}$$

is positive on the domain of the unknown parameter $p_\cdot$ of $E_\cdot$, and we easily get unimodality on the domain, hence restriction 4 was proved.

Finally we shall show restriction 5. Now we can observe the normalization of

$$\prod_{i=1}^{n_\cdot} p_\cdot^{x_\cdot}(1-p_\cdot)^{1-x_\cdot} \tag{3.9}$$

has following normalizing constant

$$\frac{1}{\int_0^1 \prod_{i=1}^{n_\cdot} p_\cdot^{x_\cdot}(1-p_\cdot)^{1-x_\cdot}dp_\cdot} \tag{3.10}$$

If we denote $m_\cdot=\sum_{i=1}^{n_\cdot}x_i$ under fixed $E_\cdot$ then the constant value equals to,

$$\frac{1}{\int_0^1 p.^{m.}(1-p.)^{n.-m.}dp.}=\frac{1}{B(m.+1,\ n.-m.+1)}$$

(3.11)

$$=\frac{\Gamma(n.+2)}{\Gamma(m.+1)\Gamma(n.-m.+1)}=\frac{(n.+1)\,!}{n.\,!\,(n.-m.)\,!}\,.$$

Next, under $n.$ observations of fixed trial $E.$ if we assume $p.$ is fixed and sample size $n.$ of $E.$ increases to infinite then the normalization

$$\prod_{i=1}^{n.} p.^{x.}(1-p.)^{1-x.}=(n.+1)\binom{n.}{m.}p.^{m.}(1-p.)^{n.-m.}$$

approaches to the function $(n.+1)N_{m.}(n.p.,\ n.p.q.)$ and we have

$$(n.+1)N_{m.}(n.p.,\ n.p.q.)=\frac{(n.+1)}{\sqrt{2\pi n.p.q.}}\exp\left\{-\frac{(m.-n.p.)^2}{2n.p.q.}\right\}$$

(3.12)

$$\sim\frac{1}{\sqrt{2\pi}\sqrt{\frac{p.q.}{n.}}}\exp\left\{-\frac{\left(\frac{m.}{n.}-p.\right)^2}{2\sqrt{\frac{p.q.}{n.}}}\right\}$$

$$=N_{p.}\left(\frac{m.}{n.},\ \frac{p.q.}{n.}\right)\sim N_{p.}\left(\frac{m.}{n.},\ \frac{m.(n.-m.)}{n.^3}\right).$$

Then we can observe the limit function of a posteriori density function

(3.13)
$$(n.+1)\binom{n.}{m.}p.^{m.}(1-p.)^{n.-m.}$$

converges to an one point concentlated density function, that is, a density function concentlated on our unknown parameter $p.$ with probability one discretely. Hence we can get restriction 5 for this binomial model as to be proved.

Acknowledgement

I am deeply indebted to Prof. Dr. K. Kunisawa in preparing this paper.

References

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

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

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

DEPARTMENT OF MATHEMATICS,
TOKYO INSTITUTE OF TECHNOLOGY.