KÔDAI MATH. SEM. REP.
22 (1970), 251–271

THE ASYMPTOTIC DISTRIBUTION OF INFORMATION PER UNIT COST CONCERNING A LINEAR HYPOTHESIS FOR MEANS OF GIVEN TWO NORMAL POPULATIONS

BY KAZUTOMO KAWAMURA

§ 1. Introduction.

On the sequential design problem Chernoff [1] has studied a sequential testing problem concerned with composit hypothesis. In his paper he has shown an essential and simple example in which he treated two mutually independent Bernoulli trials T_1 and T_2. If we denote the probability of success of trial T_1 with p_1 and the probability of success of the trial T_2 with p_2 and the hypothesis $p_1=p_2$ with H_0 and $p_1 \neq p_2$ with H_1, the subject which he treated was sequential test of the hypothesis H_0. He has given a selecting way of trial at each step definitely considering the results of preceding observations. More precisely, the procedure is deterministically given step by step comparing with Kullback-Leibler (K-L) informations of T_1 and T_2.

We have studied in [2] the asymptotic behavior of the sum of informations which discriminate the hypotheses H_0 and H_1 gained between first and n-th step under the procedure using K-L information deterministically at each step.

In our paper [4], for given finite number of populations E_i $(i=1, \cdots, k)$ which has a distribution of exponential type with one dimensional parameter θ_i $(i=1, \cdots, k)$ respectively, we had treated the sequential testing problem with respect to the given hypothesis $\mu \cdot \theta = p$, concerning unknown parameters $\theta_1, \cdots, \theta_k$, in k dimensional $(\theta_1, \cdots, \theta_k)$ space. The distribution of i-th population E_i $(i=1, \cdots, k)$ was restricted by an exponential type introduced by S. Kullback. We have given in [2], [3] a cost optimal procedure $\mathcal{P}$ selecting the populations and in [4] the equivalent randomized procedure $\mathcal{P}^*$ the limiting property of the logarithm of the likelihood ratio per unit cost concerning the hypothesis $\mu \cdot \theta = p$ of our unknown k dimensional parameter $\theta = (\theta_1, \cdots, \theta_k)$.

We have specially had some interests on the asymptotic property of the deterministic procedure. Given two trials T_1 and T_2 each of which has a normal distribution with mean m_1 and m_2 and variance σ_1^2 and σ_2^2 respectively, then the hypothesis H_0 becomes $m_1=m_2$ and H_1 becomes $m_1 \neq m_2$ analogously. In this model using the deterministic procedure which compares with K-L informations of the given trials T_1 and T_2, the selecting ratio of T_1 and T_2 has strait convergence property to the optimal ratio as given in [3] and [4] as "special example".

Received December 27, 1969.

In this paper we treat two trials E_1 and E_2 each of which has normal distribution with unknown mean m_1 and m_2 and known variance σ_1^2 and σ_2^2 respectively. Under the deterministic procedure using K-L informations of E_1 and E_2 we shall show the asymptotic distribution of the sum of informations between the first step and n-th step discriminating two hypotheses H_0 and H_1. The expected value and the order of the variance of the sum of informations will be obtained in the following sections.

In section 2 of this paper we assume that we have given a normal population $N(m, \sigma^2)$ and its n independent samples $x_1, \cdots, x_n$, we shall show the asymptotic behavior of the sum of self informations of the unknown mean m, the logarithm of the likelihood of m and the logarithm of the maximum likelihood of m. In section 3, 4 we shall restrict the number of populations to 2. And the distributions of E_1, E_2 are normal with means m_1, m_2 and variances σ_1^2, σ_2^2 where m_1, m_2 are unknown values. In section 4 the main result of this paper will be shown, that is, under the procedure $\mathscr{P}$ we have the limiting property and the asymptotic distribution of the gained sum of self informations $s_n(x_1, \cdots, x_n)$, the logarithm of the maximum likelihood of (m_1, m_2), the logarithm of the likelihood ratio under the hypotheses $m_1=m_2$ and $m_1 \neq m_2$ and the logarithm of the likelihood ratio per unit cost.

§ 2. Asymptotic behavior of the sum of self information and the logarithm of the maximum likelihood.

2. 1. Here we consider the asymptotic behavior of the sum of self informations of n independent random variables $X_1, \cdots, X_n$ from a given normal population $N(m, \sigma^2)$. The sum of self informations $s_n(X_1, \cdots, X_n)$ of the n independent random variables $X_1, \cdots, X_n$ is given by

$$s_n(X_1, \cdots, X_n)=\sum_{i=1}^{n} \log f(X_i, m, \sigma^2), \tag{2. 1}$$

where $f(X, m, \sigma^2)$ is a normal density function of population $N(m, \sigma^2)$, i.e.

$$f(X, m, \sigma^2)=\frac{1}{\sqrt{2\pi}\,\sigma} e^{-(X-m)^2/2\sigma^2}. \tag{2. 2}$$

2. 1. 1. First we consider the distribution of $s_n(X_1, \cdots, X_n)$ for fixed n. From the equality (2. 1), (2. 2) the random variable $s_n(X_1, \cdots, X_n)$ becomes

$$\begin{aligned} s_n(X_1, \cdots, X_n) &= \sum_{i=1}^{n} \log \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(X_i-m)^2/2\sigma^2} \\ &= -\frac{1}{2}\sum_{i=1}^{n}\left(\frac{X_i-m}{\sigma}\right)^2 - n\log\sqrt{2\pi}\,\sigma, \end{aligned} \tag{2. 3}$$

where $(X_i-m)/\sigma$ is a normal random variable with meae zero, variance one. Since $X_1, \cdots, X_n$ are independent, then $(X_1-m)/\sigma, \cdots, (X_n-m)/\sigma$ are independent random variables. In the following lines we put

$$\chi_n^2=\left(\frac{X_1-m}{\sigma}\right)^2+\cdots+\left(\frac{X_n-m}{\sigma}\right)^2, \tag{2. 4}$$

then χ_n^2 is distributed as χ^2-distribution with n degree of freedom. Therefore $s_n(X_1, \cdots, X_n)$ is given by a linear form of the random variable χ_n^2 as given in (2. 3)

$$s_n(X_1, \cdots, X_n)=-\frac{1}{2}\chi_n^2-n\log\sqrt{2\pi}\,\sigma, \tag{2. 5}$$

where the mean value of $s_n(X_1, \cdots, X_n)$ is given by

$$\begin{aligned} Es_n(X_1, \cdots, X_n)&=-\frac{1}{2}E\chi_n^2-n\log\sqrt{2\pi}\,\sigma \\ &=-\frac{n}{2}-n\log\sqrt{2\pi}\,\sigma, \end{aligned}$$

and the variance of $s_n(X_1, \cdots, X_n)$ is given by

$$\mathrm{Var}\, s_n(X_1, \cdots, X_n)=\left(-\frac{1}{2}\right)^2\mathrm{Var}\,(\chi_n^2)=\frac{1}{2}\cdot 2n=\frac{n}{2}.$$

2. 1. 2. Next we treat the normal approximation of the distribution of $s_n(X_1, \cdots, X_n)/n$.

$$\frac{s_n(X_1, \cdots, X_n)}{n}=-\frac{1}{2}\frac{\chi_n^2}{n}-\log\sqrt{2\pi}\,\sigma.$$

First we consider the distribution of

$$\frac{\chi_n^2}{n}=\frac{1}{n}\left(\left(\frac{X_1-m}{\sigma}\right)^2+\cdots+\left(\frac{X_n-m}{\sigma}\right)^2\right).$$

By the central limit theorem χ_n^2/n is asymptotically normally distributed with mean one and variance $2/n$ in the sense of convergence in distibution. So that $(-1/2)\chi_n^2/n$ is asymptotically normally distributed with mean $-1/2-(1/2)\log 2\pi\sigma^2$ and variance $1/2n$.

2. 1. 3. Finally we consider the limit value of a random variable $s_n(X_1, \cdots, X_n)/n$. In the equation (2. 5) χ_n^2/n is the sum of square of n independent random variables with common density $N(0, 1)$

$$\left(\frac{X_1-m}{\sigma}\right), \cdots, \left(\frac{X_n-m}{\sigma}\right).$$

Then by the strong law of large numbers we get

$$P\left\{\lim_{n\to\infty}\frac{1}{n}\left(\left(\frac{X_1-m}{\sigma}\right)^2+\cdots+\left(\frac{X_n-m}{\sigma}\right)^2\right)=E\left(\frac{X-m}{\sigma}\right)^2=1\right\}=1. \tag{2.6}$$

Therefore we have

$$\lim_{n\to\infty}\frac{s_n(X_1,\cdots,X_n)}{n}=-\frac{1}{2}-\log\sqrt{2\pi}\,\sigma \tag{2.7}$$

with probability one.

2.2.4. In this place we consider the asymptotic behavior of the maximum likelihood given by n independent random variables $X_1,\cdots,X_n$ from the population given in the preceding section. And we assume that the mean value is unknown. We define the logarithm of the likelihood function of m by

$$\log\prod_{i=1}^{n}f(X_i,m,\sigma^2)=\sum_{i=1}^{n}\log\frac{1}{\sqrt{2\pi}\,\sigma}e^{-(X_i-m)^2/2\sigma^2}. \tag{2.8}$$

Then the logarithm of the likelihood function becomes our sum of self informations $s_n(X_1,\cdots,X_n)$ defined in preceding section. The maximum likelihood of the parameter m is given by the maximum value of the likelihood function of m with respect to m.

$$\max_m\log\prod_{i=1}^{n}f(X_i,m,\sigma^2)=\max_m\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{X_i-m}{\sigma}\right)^2-n\log\sqrt{2\pi}\,\sigma\right\}. \tag{2.9}$$

The maximum value is given by differentiation of the logarithm of the likelihood function with respect to m. The value of the maximum estimate of m is given by

$$\hat{m}_n=\frac{\sum_{i=1}^{n}X_i}{n}. \tag{2.10}$$

Then the maximum value becomes

$$\begin{aligned}\log\prod_{i=1}^{n}f(X_i,\hat{m}_n,\sigma^2)&=-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{X_i-\hat{m}_n}{\sigma}\right)^2-n\log\sqrt{2\pi}\,\sigma\\&=-\frac{n}{2\sigma^2}\left(\frac{\sum_{i=1}^{n}X_i^2}{n}-\hat{m}_n{}^2\right)-n\log\sqrt{2\pi}\,\sigma,\end{aligned} \tag{2.11}$$

where $\sum_{i=1}^{n}((X_i-\hat{m}_n)/\sigma)^2$ is distributed as χ^2 distribution with freedom $n-1$. Therefore we put it as $\chi_{n-1}{}^2$. Then

$$\max_m\log\prod_{i=1}^{n}f(X_i,m,\sigma^2)=-\frac{1}{2}\chi_{n-1}^2-n\log\sqrt{2\pi}\,\sigma, \tag{2.12}$$

where the expected value of the logarithm of the maximum likelihood is

$$
\begin{aligned}
E \log \prod_{i=1}^{n} f(X_i, \hat{m}_n, \sigma^2) &= -\frac{1}{2} E\chi_{n-1}{}^2 - n \log \sqrt{2\pi}\,\sigma \\
&= -\frac{n-1}{2} - n \log \sqrt{2\pi}\,\sigma
\end{aligned} \tag{2.13}
$$

and the variance is given by

$$
\begin{aligned}
\operatorname{Var} \log \prod_{i=1}^{n} f(X_i, \hat{m}_n, \sigma^2) &= \left(-\frac{1}{2}\right)^2 \operatorname{Var}(\chi_{n-1}{}^2) \\
&= \frac{1}{4} \cdot 2 \cdot (n-1) = \frac{n-1}{2}.
\end{aligned} \tag{2.14}
$$

As a conclusion the logarithm of the maximum likelihood of m given in (2.12) is distributed as a linear form of a χ^2 random variable $\chi_{n-1}{}^2$ with $n-1$ degree of freedom.

Next we consider the asymptotic behavior of $\max_m s_n(X_1, \cdots, X_n)/n$.

$$
\begin{aligned}
\max_m s_n(X_1, \cdots, X_n)/n &= \frac{1}{n} \max_m \sum_{i=1}^{n} \log f(X_i, m, \sigma^2) \\
&= \frac{1}{n} (\log f(X_1, \hat{m}_n, \sigma^2) + \cdots + \log f(X_n, \hat{m}_n, \sigma^2)) \\
&= -\frac{1}{2} \cdot \frac{n-1}{n} \cdot \frac{X_{n-1}{}^2}{n-1} - \log \sqrt{2\pi}\,\sigma.
\end{aligned} \tag{2.15}
$$

By the central limit theorem $\chi_{n-1}{}^2/(n-1)$ is asymptotically normally distributed with mean one and variance $2/(n-1)$ in the sense of convergence in distribution. Therefore $(-1/2)\ \chi_{n-1}{}^2/n$ is asymptotically distributed with mean $-(n-1)/2n$ and variance $(n-1)/2n^2$. Then the logarithm of the maximum likelihood per unit sample $\max_m \sum_{i=1}^{n} \log f(X_i, m, \sigma^2)/n$ is asymptotically distributed as normal distribution with mean $-(n-1)/2n - \log \sqrt{2\pi}\sigma$ and variance $(n-1)/2n^2$.

Finally we consider the limit value as $n \to \infty$ of the logarithm of the maximum likelihood:

$$
\begin{aligned}
\max_m \log \prod_{i=1}^{n} f(X_i, m, \sigma^2) &= -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \hat{m}_n)^2 - n \log \sqrt{2\pi}\,\sigma \\
&= -\frac{n}{2\sigma^2} \left(\frac{1}{n} \sum_{i=1}^{n} X_i{}^2 - \hat{m}_n{}^2\right) - n \log \sqrt{2\pi}\,\sigma.
\end{aligned} \tag{2.11}
$$

And, by the strong law of large numbers, $\sum_{i=1}^{n} X_i{}^2/n \to E(X^2)$ as $n \to \infty$ with probability one. Then the logarithm of the maximum likelihood per unit sample has a limit value as followings.

$$\lim_{n\to\infty}\frac{1}{n}\max_m \log\prod_{i=1}^{n} f(X_i, m, \sigma^2)$$

$$(2.16)\qquad = -\frac{1}{2\sigma^2}\{E(X^2)-E^2(X)\} - \log\sqrt{2\pi}\,\sigma$$

$$= -\frac{1}{2\sigma^2}\{\sigma^2\} - \log\sqrt{2\pi}\,\sigma = -\frac{1}{2} - \log\sqrt{2\pi}\,\sigma$$

with probability one.

In this section we have considered the asymptotic behavior of the sum of self informations $s_n(X_1, \cdots, X_n)$ given by $X_1, \cdots, X_n$ and the logarithm of the maximum likelihood $\max_m s_n(X_1, \cdots, X_n)$ with respect to unknown mean m for n independent random variables $X_1, \cdots, X_n$ from a given normal population $N(m, \sigma^2)$.

NOTE 1. For n independent random valiables $X_1, \cdots, X_n$ from $N(m, \sigma^2)$, the difference between the two values $\max_m \log \Pi_{i=1}^n f(X_i, m, \sigma^2)$ and $s_n(X_1, \cdots, X_n)$

$$\max_m \log\prod_{i=1}^{n} f(X_i, m, \sigma^2) - s_n(X_1, \cdots, X_n)$$

$$= \left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 - n\log\sqrt{2\pi}\,\sigma\right\} - \left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{X_i-m}{\sigma}\right)^2 - n\log\sqrt{2\pi}\,\sigma\right\}$$

$$(2.17)\qquad = \frac{1}{2\sigma^2}\sum_{i=1}^{n}\{(X_i-m)^2-(X_i-\bar{X})^2\}$$

$$= \frac{1}{2\sigma^2}\{nm^2 - 2nm\bar{X} + n\bar{X}^2\} = \frac{1}{2}\left(\frac{\bar{X}-m}{\sigma/\sqrt{n}}\right)^2$$

is distributed as χ^2 distribution with one degree of freedom.

NOTE 2. It holds the next relation:

$$(2.18)\qquad \frac{\max_m s_n(X_1, \cdots, X_n)}{n} - \frac{s_n(X_1, \cdots, X_n)}{n} = \frac{1}{2n}\chi_1^2.$$

Therefore the difference has mean $1/2n$ and variance $1/2n^2$.

NOTE 3. If we have two independent samples $X_1, \cdots, X_n$ and $X_1^*, \cdots, X_n^*$ with size n from the normal population $N(m, \sigma^2)$, then $s_n(X_1, \cdots, X_n)/n$ is asymptotically distributed as normal distribution $N(-1/2-\log\sqrt{2\pi}\,\sigma, 1/2n)$ and $\max_m s_n(X_1^*, \cdots, X_n^*)/n$ is asymptotically distribued as normal distribution $N(-(n-1)/2n-\log\sqrt{2\pi}\,\sigma, (n-1)/2n^2)$. Therefore the difference $\max_m s_n(X_1^*, \cdots, X_n^*)/n - s_n(X_1, \cdots, X_n)/n$ is asymptotically normally distributed with mean

$$(2.19)\qquad \left(-\frac{n-1}{2n} - \log\sqrt{2\pi}\,\sigma\right) - \left(-\frac{1}{2} - \log\sqrt{2\pi}\,\sigma\right) = \frac{1}{2n}$$

and variance

$$\frac{n-1}{2n^2}+\frac{1}{2n}=\frac{1}{n}-\frac{1}{2n^2} \tag{2. 20}$$

that is, the difference is asymptotically considered as a random variable of normal population $N(1/2n, 1/n-1/2n^2)$.

§ 3. Most informative procedure concerning the costs of experiments.

In this section we shall discuss the procedure of selection of two normal populations $N(m_1, \sigma_1^2)$, $N(m_2, \sigma_2^2)$, where the means m_1, m_2 are unknown and the variances are known. And if we select $N(m_i, \sigma_i^2)$ $(i=1, 2)$ we must pay cost C_i $(i=1, 2)$ for each sample from the populations.

First we define the aim of policy of selecting populations. Under the given aim we shall define the optimal policy $\mathscr{P}$ of selecting populations in each steps. And we shall define equivalent, in limiting property, randomized policy of selecting populations $\mathscr{P}^*$. Under the policy $\mathscr{P}$, $\mathscr{P}^*$ we have studied the limiting optimality and the equivalence in limiting property of the two policy $\mathscr{P}$, $\mathscr{P}^*$ in [3].

In 3. 1 we shall discuss, under the sequential deciding procedure $\mathscr{P}$, the asymptotic behavior of the objective function as given in [2], [3] etc., specially the logarithm of the likelihood ratio per unit cost $S_n(\hat{\theta}_n, \tilde{\theta}_n)/\sum_{i=1}^n C^{(i)}$.

3. 1. Definition of the optimal procedure.

The aim of our procedure of selecting the two trials E_1 and E_2 is to discriminate whether the unknown pair of means m_1 and m_2 exists on a given linear line $m_1=m_2$ or not. To discriminate the two hypotheses $m_1=m_2$ and $m_1\neq m_2$ we additionally consider the optimality in the sense of discrimination per unit cost.[1)] We considered to pay costs C_1 or C_2 to observe samples from the two trials E_1 or E_2 respectively. For the discrimination we define the logarithm of the likelihood ratio $S_n(\hat{\theta}_n, \tilde{\theta}_n)$ as followings:

$$S_n(\hat{\theta}_n, \tilde{\theta}_n)=\frac{\max\limits_{\theta\in R^2} \log \prod\limits_{i=1}^n f(X_i, \theta, E^{(i)})}{\max\limits_{\theta\in A(\hat{\theta}_n)} \log \prod\limits_{i=1}^n f(X_i, \theta, E^{(i)})}, \tag{3. 1}$$

where $X_1, \cdots, X_n$ are the first n observations, $E^{(i)}$ $(i=1, \cdots, n)$ means the i-th selected trial, θ means the two dimensional mean (m_1, m_2) of the trials E_1, E_2 and $\theta=\hat{\theta}_n$ gives the logarithm of the likelihood

$$\max_{\theta\in R^2} \log \prod_{i=1}^n f(X_i, \theta, E^{(i)})=\log \prod_{i=1}^n f(X_i, \hat{\theta}_n, E^{(i)}). \tag{3. 2}$$

In the following lines we define $A(\hat{\theta}_n)$ as the alternative domain of $\hat{\theta}_n$ in the 2-

1) See [2] and the generalized form [3].

dimensional Euclidean space R^2. On the alternative domain there exists unique $\tilde{\theta}_n$ which maximizes the likelihood function

$$\max_{\theta\in A(\hat{\theta}_n)} \log \prod_{i=1}^{n} f(X_i, \theta, E^{(i)}) = \log \prod_{i=1}^{n} f(X_i, \tilde{\theta}_n, E^{(i)}). \tag{3. 3}$$

We put n_1 as the number of E_1 in the first n selections $E^{(1)}, \cdots, E^{(n)}$ and we define n_2 so that $n_1+n_2=n$. Then the unique $\hat{\theta}_n$ is given by

$$\hat{\theta}_n = \left(\frac{\sum_{E^{(i)}=E_1} X_i}{n_1}, \frac{\sum_{E^{(i)}=E_2} X_i}{n_2}\right). \tag{3. 4}$$

And $\tilde{\theta}_n$ is determined uniquely as following way. Put

$$\begin{aligned}
L(\theta) &= \log \prod_{i=1}^{n} f(X_i, \theta, E^{(i)}) \\
&= \log \prod_{E^{(i)}=E_1} f(X_i, \theta, E^{(i)}) + \log \prod_{E^{(i)}=E_2} f(X_i, \theta, E^{(i)}) \\
&= \sum_{E^{(i)}=E_1} \log \frac{1}{\sqrt{2\pi}\,\sigma_1} e^{-(X_i-m_1)^2/2\sigma_1^2} + \sum_{E^{(i)}=E_2} \log \frac{1}{\sqrt{2\pi}\,\sigma_2} e^{-(X_i-m_2)^2/2\sigma_2^2} \\
&= \left(n_1 \log \frac{1}{\sqrt{2\pi}\,\sigma_1} + n_2 \log \frac{1}{\sqrt{2\pi}\,\sigma_2}\right) - \sum_{E^{(i)}=E_1} \frac{(X_i-m_1)^2}{2\sigma_1^2} - \sum_{E^{(i)}=E_2} \frac{(X_i-m_2)^2}{2\sigma_2^2}.
\end{aligned}$$

Then we have easily that $\max_{\theta\in A(\hat{\theta}_n)} L(\theta)$ is ginen uniquely on the boundary π of hypotheses $m_1=m_2$, so that we have

$$\max_{\theta\in A(\hat{\theta}_n)} L(\theta) = \max_{\theta\in\pi} L(\theta). \tag{3. 5}$$

Therefore, $\tilde{\theta}_n$ is an element of π. If we put $m_1=m_2=m$, then

$$\max_{\theta\in A(\hat{\theta}_n)} L(\theta) = \max_{m} L(\theta),$$

$$\begin{aligned}
\frac{\partial L}{\partial m} &= \frac{\partial}{\partial m}\left(-\frac{\sum_{E^{(i)}=E_1}(X_i-m)^2}{2\sigma_1^2} - \frac{\sum_{E^{(i)}=E_2}(X_i-m)^2}{2\sigma_2^2}\right) \\
&= \frac{\sum_{E^{(i)}=E_1}(X_i-m)}{2\sigma_1^2} + \frac{\sum_{E^{(i)}=E_2}(X_i-m)}{2\sigma_2^2} \\
&= \frac{\sum_{E^{(i)}=E_1} X_i}{\sigma_1^2} + \frac{\sum_{E^{(i)}=E_2} X_i}{\sigma_2^2} - \left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)m.
\end{aligned}$$

If we put $\partial L/\partial m=0$ then $m=(\sum_{E_1} X_i/\sigma_1^2 + \sum_{E_2} X_i/\sigma_2^2)/(n_1/\sigma_1^2+n_2/\sigma_2^2)$. Then we have

$$\tilde{\theta}_n = \left(\left(\frac{\sum_{E_1} X_i}{\sigma_1^2} + \frac{\sum_{E_2} X_i}{\sigma_2^2}\right)\Big/\left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right), \left(\frac{\sum_{E_1} X_i}{\sigma_1^2} + \frac{\sum_{E_2} X_i}{\sigma_2^2}\right)\Big/\left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)\right)$$

$$\tilde{\theta}_n=\left(\left(\frac{n_1}{\sigma_1^2}\frac{\sum_{E_1}X_i}{n_1}+\frac{n_2}{\sigma_2^2}\frac{\sum_{E_2}X_i}{n_2}\right)\Big/\left(\frac{n_1}{\sigma_1^2}+\frac{n_2}{\sigma_2^2}\right),\right.$$
$$\left.\left(\frac{n_1}{\sigma_1^2}\frac{\sum_{E_1}X_i}{n_1}+\frac{n_2}{\sigma_2^2}\frac{\sum_{E_2}X_i}{n_2}\right)\Big/\left(\frac{n_1}{\sigma_1^2}+\frac{n_2}{\sigma_2^2}\right)\right),$$

where $\sum_{E_1}$ means the sum of i such that $E^{(i)}=E_1$ in the first n trials $E^{(1)},\cdots,E^{(n)}$ and also $\sum_{E_2}$ means the sum of i such that $E^{(i)}=E_2$ in the first n trials. And we put $r_n=n_1/\sigma_1^2/(n_1/\sigma_1^2+n_2/\sigma_2^2)$, then $1-r_n=n_2/\sigma_2^2/(n_1/\sigma_1^2+n_2/\sigma_2^2)$,

$$\tilde{\theta}_n=\left\{r_n\frac{\sum_{E_1}X_i}{n_1}+(1-r_n)\frac{\sum_{E_2}X_i}{n_2},\ r_n\frac{\sum_{E_1}X_i}{n_1}+(1-r_n)\frac{\sum_{E_2}X_i}{n_2}\right\}. \tag{3.6}$$

And (3. 1) becomes

$$S_n(\hat{\theta}_n,\tilde{\theta}_n)=\sum_{i=1}^{n}\log\frac{f(X_i,\hat{\theta}_n,E^{(i)})}{f(X_i,\tilde{\theta}_n,E^{(i)})}. \tag{3.7}$$

Next we introduce the notion of costs, that is, for any step we pay the costs C_j for the observation from the trials E_j $(j=1,2)$. The sum of costs between the first and n-th step is $\sum_{i=1}^{n}C^{(i)}$, where $C^{(1)},\cdots,C^{(n)}$ are the sequence of costs paid for these steps.

DEFINITION OF THE PROCEDURE $\mathcal{P}$. For any step n the ratio

$$\frac{S_n(\hat{\theta}_n,\tilde{\theta}_n)}{\sum_{i=1}^{n}C^{(i)}} \tag{3.8}$$

is considered as a random variable. The limiting property of the random variable as $n\to\infty$ for some selecting way of $E^{(1)},E^{(2)},\cdots$ is given in [2], [3] and [4]. We shall call a selecting way of $E^{(1)},E^{(2)},\cdots$ as a procedure (or policy) in the following lines. Under any procedure $\mathcal{P}$ having selecting ratio λ of E_1 the sequence of the random variables (3. 8) has a limiting value. So we may classify the procedure by the limiting ratio λ of selecting E_1 and put the class of the procedure as $\mathcal{P}_\lambda$. By an easy calculation the right side of equality (3. 1) becomes

$$S_n(\hat{\theta}_n,\tilde{\theta}_n)=n_1I(\hat{\theta}_n,\tilde{\theta}_n,E_1)+n_2I(\hat{\theta}_n,\tilde{\theta}_n,E_2) \tag{3.9}$$

where $I(\theta,\varphi,E_j)$ is the mean discrimination defined by S. Kullback [5]:

$$I(\theta,\varphi,E_j)=\int\log\frac{f(X,\theta,E_j)}{f(X,\varphi,E_j)}f(X,\theta,E_j)dX. \tag{3.10}$$

The sum of costs paid in the first n steps is

$$\sum_{i=1}^{n}C^{(i)}=n_1C_1+n_2C_2.$$

Therefore, to maximize the ratio

$$\text{(3. 11)} \qquad \frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}} = \frac{n_1 I(\hat{\theta}_n, \tilde{\theta}_n, E_1) + n_2 I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{n_1 C_1 + n_2 C_2}$$

we may define the procedure $\mathscr{P}$ as $n+1$-th step $E^{(n+1)}$ as follows:

$$\text{(3. 12)} \qquad E^{(n+1)} = \begin{Bmatrix} E_1 \\ E_2 \\ E^{(n)} \end{Bmatrix} \quad \text{if} \quad \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_1)}{C_1} \begin{Bmatrix} > \\ < \\ = \end{Bmatrix} \frac{I(\hat{\theta}_n, \tilde{\theta}_n, E_2)}{C_2}.$$

In fact we can verify the Kullback's mean information for discrimination (3. 10) as followings:

$$\text{(3. 13)} \qquad I(\theta, \varphi, E_j) = \frac{(m_j - m_j^*)^2}{2\sigma_j^2} \qquad (j=1, 2),$$

where $\theta=(m_1, m_2)$, $\varphi=(m_1^*, m_2^*)$. From $\hat{\theta}_n$ in (3. 4), $\tilde{\theta}_n$ in (3. 6) and the equality (3. 13) our procedure (3. 12) becomes

$$\text{(3. 14)} \qquad \begin{aligned} E^{(n+1)} = \begin{Bmatrix} E_1 \\ E_2 \\ E^{(n)} \end{Bmatrix} \quad \text{if} \quad & \frac{1}{2\sigma_1^2 C_1}\left[\frac{\sum_{E_1} X_i}{n_1} - \left\{r_n \frac{\sum_{E_1} X_i}{n_1} + (1-r_n)\frac{\sum_{E_2} X_i}{n_2}\right\}\right]^2 \\ & \begin{Bmatrix} > \\ < \\ = \end{Bmatrix} \frac{1}{2\sigma_2^2 C_2}\left[\frac{\sum_{E_2} X_i}{n_2} - \left\{r_n \frac{\sum_{E_1} X_i}{n_1} + (1-r_n)\frac{\sum_{E_2} X_i}{n_2}\right\}\right]^2 \end{aligned}$$

or equivalently

$$\text{(3. 15)} \quad \text{if} \quad \frac{1}{2\sigma_1^2 C_1}(1-r_n)^2\left(\frac{\sum_{E_1} X_i}{n_1} - \frac{\sum_{E_2} X_i}{n_2}\right)^2 \begin{Bmatrix} > \\ < \\ = \end{Bmatrix} \frac{1}{2\sigma_2^2 C_2} r_n^2 \left(\frac{\sum_{E_1} X_i}{n_1} - \frac{\sum_{E_2} X_i}{n_2}\right)^2$$

or equivalently with probability one

$$\text{(3. 16)} \qquad \text{if} \quad \frac{(1-r_n)^2}{\sigma_1^2 C_1} \begin{Bmatrix} > \\ < \\ = \end{Bmatrix} \frac{r_n^2}{\sigma_2^2 C_2}.$$

And by definition of r_n used in (3. 6) we have

$$\frac{1-r_n}{r_n} = \left\{\frac{n_2}{\sigma_2^2}\Big/\left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)\right\}\Big/\left\{\frac{n_1}{n_1^2}\Big/\left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)\right\} = \frac{n_2}{\sigma_2^2}\Big/\frac{n_1}{\sigma_1^2}.$$

Then (3. 16) becomes equivalently

$$\text{(3. 17)} \qquad \text{if} \quad \left(\frac{n_2}{\sigma_2^2}\Big/\frac{n_1}{\sigma_1^2}\right)^2 \begin{Bmatrix} > \\ < \\ = \end{Bmatrix} \frac{\sigma_1^2 C_1}{\sigma_2^2 C_2}$$

or equivalently

$$\text{if}\quad \frac{n_2}{n_1}\begin{Bmatrix} > \\ < \\ = \end{Bmatrix}\frac{\sigma_2\sqrt{C_1}}{\sigma_1\sqrt{C_2}} \tag{3.18}$$

or equivalently

$$\text{if}\quad \frac{n}{n_1}=\frac{n_2}{n_1}+1\begin{Bmatrix} > \\ < \\ = \end{Bmatrix}\frac{\sigma_2\sqrt{C_1}+\sigma_1\sqrt{C_2}}{\sigma_1\sqrt{C_2}}.$$

Therefore we get equivalently

$$\text{if}\quad \frac{n_1}{n}\begin{Bmatrix} < \\ > \\ = \end{Bmatrix}\frac{\sigma_1\sqrt{C_2}}{\sigma_2\sqrt{C_1}+\sigma_1\sqrt{C_2}} \tag{3.19}$$

with probability one. In the following lines we put right side of (3. 19) as λ then

$$\lambda=\frac{\sigma_1\sqrt{C_2}}{\sigma_2\sqrt{C_1}+\sigma_1\sqrt{C_2}}. \tag{3.20}$$

LEMMA 1. *For some procedure if* $\min(n_1, n_2)\to\infty$ *as* $n\to\infty$, *then* $\hat{\theta}_n$ *converges to the pair of unknown parameters* (m_1, m_2) *with probability one.*

LEMMA 2. *For some procedure if* $n_1/n\to\lambda$ *as* $n\to\infty$, *then* $\tilde{\theta}_n$ *converges to* θ^* *with probabiliy one as* $n\to\infty$. *Where* θ^* *is given by the equations*

$$\frac{(m_1-m_1^*)^2}{2C_1\sigma_1^2}=\frac{(m_2-m_2^*)^2}{2C_2\sigma_2^2}\qquad \textit{and}\quad m_1^*=m_2^*. \tag{3.21}$$

In the paper [4] the conditions in lemma 1, lemma 2:

I. $\min(n_1, n_2)\to\infty$ as $n\to\infty$ with probability one,

II. $n_1/n\to\lambda$ as $n\to\infty$ with probability one

are called as the optimal conditions and we have shown in [4] that under the procedure $\mathcal{P}$ we can get the property conditions I, II.

LEMMA 3. *The pocedure* $\mathcal{P}$ *has the property of optimal conditions.*

Therefore under the lemma 1～3 we have

THEOREM 1. *Under the procedure* $\mathcal{P}$ *we have*

$$\lim_{n\to\infty}\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=I^*(\theta) \tag{3.22}$$

with probability one as given in [2]. *The value* $I^*(\theta)$ *is given by*

$$\frac{I(\theta, \theta^*, E_1)}{C_1}=\frac{I(\theta, \theta^*, E_2)}{C_2}=I^*(\theta) \tag{3.23}$$

where θ is the unknown parameter and θ^* is the root of the equation $m_1=m_2$ given in (3. 23) *or see* (3. 21).

THEOREM 2. *Under the other procedure $\mathcal{P}'$ if the ratio $S_n(\hat{\theta}_n, \tilde{\theta}_n)/\sum_{i=1}^n C^{(i)}$ converges to the limit value $I^{**}(\theta)$, then*

$$I^{**}(\theta) \leqq I^*(\theta). \tag{3. 24}$$

The proof is given in [4] so that our procedure $\mathcal{P}$ is the asymptotically optimal procedure in the sense described above.

3. 2. The randomized optimal procedure.

In preceding 3. 1 we have observed that the procedure $\mathcal{P}$ (3. 12) has the optimal conditions I, II and the maximum limit property in (3. 22), (3. 24). And the optimal procedure (3. 12) becomes equivalently

$$E^{(n+1)} = \begin{Bmatrix} E_1 \\ E_2 \\ E^{(n)} \end{Bmatrix} \quad \text{if} \quad \frac{n_1}{n} \begin{Bmatrix} < \\ > \\ = \end{Bmatrix} \lambda \tag{3. 25}$$

with probability one as in (3. 19), (3. 20). In this place we generalize the procedure $\mathcal{P}$ as the following binomially randomized way. We define the procedure $\mathcal{P}^*$ by $P\{E^{(n+1)}=E_1\}=\lambda$.

LEMMA. The procedure $\mathcal{P}^*$ has the optimal condition I and II.

Proof. By the strong law of large numbers $\mathcal{P}^*$ has the property that the selecting ratio of E_1 between first and n-th step n_1/n converges to the ratio λ with probability one. Therefore $\mathcal{P}^*$ has the property of optimal condition II, and by the inequality $0<\lambda<1$ we easily have $\min(n_1, n_2)\to\infty$ as $n\to\infty$ with probability one, that is, $\mathcal{P}^*$ has the optimal condition I as to be proved.

Therefore, under $\mathcal{P}^*$, we can verify the same result of lemmas 1, 2. Therefore under $\mathcal{P}^*$, we have the same result of theorem 1, 2. So that $\mathcal{P}^*$ also is an element of the optimal procedure in the sense (3. 22), (3. 24) in theorems 1, 2.

NOTE. If we put the costs C_1, C_2 of E_1, E_2 as $C_1=C_2=1$, then we have $\sum_{i=1}^n C^{(i)}=n$. Therefore, the ratio of our interest (3. 8) becomes

$$\frac{1}{n}\sum_{i=1}^{n} \log \frac{f(X_i, \hat{\theta}_n, E^{(i)})}{f(X_i, \tilde{\theta}_n, E^{(i)})} \tag{3. 26}$$

which is ordinary sample mean of self discriminations given by the n samples between first and n-th step.

§ 4. Main results.

In this section we consider the asymptotic behavior of the sum of self informations, the logarithm of the maximum likelihood, the logarithm of the likelihood ratio on the hypothesis $m_1=m_2$ and the logarithm of the likelihood ratio per unit cost under the procedure $\mathcal{P}$.

4. 1.

4. 1. 1. The asymptotic behavior of sum of self informations $s_n(X_1, \cdots, X_n)$ under $\mathcal{P}$.

$$\begin{aligned} s_n(X_1, \cdots, X_n) &= \log \prod_{i=1}^{n} f(X_i, \theta, E^{(i)}) \\ &= \log \prod_{E_1} f(X_i, \theta, E^{(i)}) + \log \prod_{E_2} f(X_i, \theta, E^{(i)}) \\ &= \log \prod_{E_1} \frac{1}{\sqrt{2\pi}\,\sigma_1} e^{-(X_i-m_1)^2/2\sigma_1^2} + \log \prod_{E_2} \frac{1}{\sqrt{2\pi}\,\sigma_2} e^{-(X_i-m_2)^2/2\sigma_2^2} \qquad (4.\ 1) \\ &= \left(n_1 \log \frac{1}{\sqrt{2\pi}\,\sigma_1} + n_2 \log \frac{1}{\sqrt{2\pi}\,\sigma_2} \right) \\ &\quad - \frac{1}{2} \left[\sum_{E_1} \left(\frac{X_i - m_1}{\sigma_1} \right)^2 + \sum_{E_2} \left(\frac{X_i - m_2}{\sigma_2} \right)^2 \right]. \end{aligned}$$

Here $\sum_{E_1}((X_i-m_1)/\sigma_1)^2$ is a random variables which is a function of n_1 independent random variables from the normal population E_1: $N(m_1, \sigma_1^2)$. And $(X_i-m_1)/\sigma_1$ is a random variables from a normal population $N(0, 1)$ if $E^{(i)}=E_1$ is satisfied, therefore $((X_i-m_1)/\sigma_1)^2$ is a random variable of χ^2 distribution with one degree of freedom, so our $\sum_{E_1}((X_i-m_1)/\sigma_1)^2$ is a random variable of χ^2 distribution with n_1 degree of freedom. And $\sum_{E_2}((X_i-m_2)/\sigma_2)^2$ is a random variable of χ^2 distribution with n_2 degree of freedom which is independent of $\sum_{E_1}((X_i-m_1)/\sigma_1)^2$. Therefore $\sum_{E_1}((X_i-m_1)/\sigma_1)^2+\sum_{E_2}((X_i-m_2)/\sigma_2)^2$ is a random variable of χ^2 distribution with $n_1+n_2=n$ degree of freedom. In the following lines we put it as χ_n^2. Then our $s_n(X_1, \cdots, X_n)$ becomes

$$s_n(X_1, \cdots, X_n) = -\frac{1}{2}\chi_n^2 + n_1 \log \frac{1}{\sqrt{2\pi}\,\sigma_1} + n_2 \log \frac{1}{\sqrt{2\pi}\,\sigma_2}. \qquad (4.\ 2)$$

4. 1. 2. The distribution of $s_n(X_1, \cdots, X_n)/n$.

Since

$$\frac{s_n(X_1, \cdots, X_n)}{n} = -\frac{1}{2}\frac{\chi_n^2}{n} + \frac{n_1}{n}\log \frac{1}{\sqrt{2\pi}\,\sigma_1} + \frac{n_2}{n}\log \frac{1}{\sqrt{2\pi}\,\sigma_2},$$

under $\mathcal{P}$, we have $|n_1/n-\lambda| \leqq 1/n$ with probability one. Therefore, by the equivalent condition of $\mathcal{P}$ (3. 25), n_1/n is a function of n with probability one. Using the result of section 2, $-(1/2n)\chi_n^2$ is asymptotically normally distributed as $N(-1/2, 1/2n)$. Then $s_n(X_1, \cdots, X_n)/n$ is asymptotically normally distributed with mean $-1/2+(n_1/n) \log(1/\sqrt{2\pi}\,\sigma_1)+(n_2/n)\log(1/\sqrt{2\pi}\,\sigma_2)$ and variance $1/2n$.

4. 2.

4. 2. 1. The asymptotic behavior of the logarithm of the maximum likelihood $\max_\theta s_n(X_1, \cdots, X_n)$.

$$\max_\theta s_n(X_1, \cdots, X_n) = \max_\theta \log \sum_{i=1}^{n} f(X_i, \theta, E^{(i)})$$

$$= \log \prod_{i=1}^{n} f(X_i, \hat{\theta}_n, E^{(i)})$$

$$(4.4) \qquad = \log \prod_{E_1} f(X_i, \hat{\theta}_n, F_1) + \log \prod_{E_2} f(X_i, \hat{\theta}_n, E_2)$$

$$= \left(n_1 \log \frac{1}{\sqrt{2\pi}\,\sigma_1} + n_2 \log \frac{1}{\sqrt{2\pi}\,\sigma_2}\right)$$

$$- \frac{1}{2} \sum_{E_1} \left(\frac{X_i - \sum_{E_1} X_i/n_1}{\sigma_1}\right)^2 - \frac{1}{2} \sum_{E_2} \left(\frac{X_i - \sum_{E_2} X_i/n_2}{\sigma_2}\right)^2;$$

where $\sum_{E_1} \{(X_i - \sum_{E_1} X_i/n_1)/\sigma_1\}^2$ is a random variable consisting of n_1 random variables from our population $N(m_1, \sigma_1^2)$ which is distributed as χ^2 distribution with $n_1 - 1$ degree of freedom, and $\sum_{E_2} \{(X_i - \sum_{E_2} X_i/n_2)/\sigma_2\}^2$ is also a random variable of χ^2 distribution with $n_2 - 1$ degree of freedom which is independent of $\sum_{E_1} \{(X_i - \sum_{E_1} X_i/n_1)/\sigma_1\}^2$. Then the sum is a random variable of χ^2 distribution with $(n_1 - 1) + (n_2 - 1) = n - 2$ degree of freedom. Therefore we get

$$(4.5) \qquad \max_\theta s_n(X_1, \cdots, X_n) = -\frac{1}{2} \chi_{n-2}{}^2 + \left(n_1 \log \frac{1}{\sqrt{2\pi}\,\sigma_1} + n_2 \log \frac{1}{\sqrt{2\pi}\,\sigma_2}\right).$$

4. 2. 2. The asymptotic behavior of $\max_\theta s_n(X_1, \cdots, X_n)/n$.

$$\frac{\max_\theta s_n(X_1, \cdots, X_n)}{n} = -\frac{1}{2} \frac{\chi_{n-2}{}^2}{n} + \frac{n_1}{n} \log \frac{1}{\sqrt{2\pi}\,\sigma_1} + \frac{n_2}{n} \log \frac{1}{\sqrt{2\pi}\,\sigma_2}$$

$$(4.6) \qquad = -\frac{1}{2} \frac{n-2}{n} \frac{\chi_{n-2}{}^2}{n-2} + \frac{n_1}{n} \log \frac{1}{\sqrt{2\pi}\,\sigma_1} + \frac{n_2}{n} \log \frac{1}{\sqrt{2\pi}\,\sigma_2}.$$

The random variable $\chi_{n-2}{}^2/(n-2)$ is asymptotically normallydistributed with mean one and variance $2/(n-2)$. Therefore $\max_\theta s_n(X_1, \cdots, X_n)/n$ is asymptotically normally distributed with mean

$$-\frac{1}{2} \frac{n-2}{n} + \frac{n_1}{n} \log \frac{1}{\sqrt{2\pi}\,\sigma_1} + \frac{n_2}{n} \log \frac{1}{\sqrt{2\pi}\,\sigma_2}$$

and variance

$$\left(-\frac{1}{2} \frac{n-2}{2}\right)^2 \frac{2}{n-2} = \left(\frac{n-2}{n}\right)^2 \frac{1}{2(n-2)} = \frac{(n-2)}{2n^2}.$$

4. 2. 3. The limit value of $\max_\theta s_n(X_1, \cdots, X_n)/n$.

In the inequality (4. 6) $\chi_{n-2}{}^2/(n-2)$ has a limit value 1 with probability one;

$P\{\lim_{n\to\infty}\chi_{n-2}^2/n-2=1\}=1$. Then we have

$$\lim_{n\to\infty}\frac{s_n(X_1,\cdots,X_n)}{n}=-\frac{1}{2}+\lambda\log\frac{1}{\sqrt{2\pi}\,\sigma_1}+(1-\lambda)\log\frac{1}{\sqrt{2\pi}\,\sigma_2} \tag{4.7}$$

with probability one.

4.3. Under the procedure $\mathcal{P}$ we consider the asymptotic distribution of $S_n(\hat\theta_n,\tilde\theta_n)/n$ and limit value.

By (3.9), (3.13) and (3.6) we have

$$\begin{aligned} S_n(\hat\theta_n,\tilde\theta_n)&=n_1I(\hat\theta_n,\tilde\theta_n,E_1)+n_2I(\hat\theta_n,\tilde\theta_n,E_2)\\ &=n_1\frac{(\sum_{E_1}X_i/n_1-m^*)^2}{2\sigma_1^2}+n_2\frac{(\sum_{E_2}X_i/n_2-m^*)^2}{2\sigma_2^2},\end{aligned} \tag{4.8}$$

where $m^*=r_n\sum_{E_1}X_i/n_1+(1-r_n)\sum_{E_2}X_i/n_2$. Then we have

$$\begin{aligned} \frac{(\sum_{E_1}X_i/n_1-m^*)^2}{2\sigma_1^2}&=\frac{(1-r_n)^2(\sum_{E_1}X_i/n_1-\sum_{E_2}X_i/n_2)^2}{2\sigma_1^2},\\ \frac{(\sum_{E_2}X_i/n_1-m^*)^2}{2\sigma_2^2}&=\frac{r_n^2(\sum_{E_1}X_i/n_1-\sum_{E_2}X_i/n_2)^2}{2\sigma_2^2}.\end{aligned} \tag{4.9}$$

Therefore we have

$$\begin{aligned} S_n(\hat\theta_n,\tilde\theta_n)&=\left\{n_1\frac{(1-r_n)^2}{2\sigma_1^2}+n_2\frac{r_n^2}{2\sigma_2^2}\right\}\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2\\ &=\frac{1}{2}\left[\frac{n_1}{\sigma_1^2}(1-r_n)^2+\frac{n_2}{\sigma_2^2}r_n^2\right]\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2,\end{aligned} \tag{4.10}$$

where $r_n=(n_1/\sigma_1^2)/(n_1/\sigma_1^2+n_2/\sigma_2^2)$. Then we get

$$\begin{aligned} S_n(\hat\theta_n,\tilde\theta_n)&=\frac{1}{2}\left\{\frac{n_1}{\sigma_1^2}\left(\frac{n_2}{\sigma_2^2}\Big/\left(\frac{n_1}{\sigma_1^2}+\frac{n_2}{\sigma_2^2}\right)\right)^2+\frac{n_2}{\sigma_2^2}\left(\frac{n_1}{\sigma_1^2}\Big/\left(\frac{n_1}{\sigma_1^2}+\frac{n_2}{\sigma_2^2}\right)\right)^2\right\}\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2\\ &=\frac{1}{2}\left\{\frac{n_1}{\sigma_1^2}\frac{n_2}{\sigma_2^2}\left(\frac{n_1}{\sigma_1^2}+\frac{n_2}{\sigma_2^2}\right)\Big/\left(\frac{n_1}{\sigma_1^2}+\frac{n_2}{\sigma_2^2}\right)^2\right\}\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2\\ &=\frac{1}{2}\left\{\frac{n_1}{\sigma_1^2}\frac{n_2}{\sigma_2^2}\Big/\left(\frac{n_1}{\sigma_1^2}+\frac{n_2}{\sigma_2^2}\right)\right\}\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2\\ &=\frac{1}{2}\frac{n_1n_2}{\sigma_2^2n_1+\sigma_1^2n_2}\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2\\ &=\frac{n}{2}\left\{\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right\}\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2\\ &=\frac{n}{2}\left\{\frac{n_1}{n}\left(1-\frac{n_1}{n}\right)\Big/\left(\frac{n_1}{n}(\sigma_2^2-\sigma_1^2)+\sigma_1^2\right)\right\}\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2.\end{aligned} \tag{4.11}$$

Following the procedure $\mathcal{P}$ for any step n the number of selections n_1 of E_1 is a fixed number depending only on n. By the argument of 3.1, of preceding section, under the procedure $\mathcal{P}$, as $(1/2)n_1n_2/(\sigma_2^2n_1+\sigma_1^2n_2)$ is a fixed number for fixed n, we see that in $S_n(\hat{\theta}_n, \tilde{\theta}_n)$ only $\sum_{E_1} X_i/n_1$, $\sum_{E_2} X_i/n_2$ are random variables. These two random variables are independent random variables which are distributed normally with means m_1, m_2 and variances σ_1^2/n_1, σ_2^2/n_2 respectively, so that $\sum_{E_1} X_i/n_1-\sum_{E_2} X_i/n_2$ is a random variable of a normal distribution with mean m_1-m_2 and variance $\sigma_1^2/n_1+\sigma_2^2/n_2$. Therefore, to know the distribution of $(\sum_{E_1} X_i/n_1-\sum_{E_2} X_i/n_2)^2$, we put $Y_n=\sum_{E_1} X_i/n_1-\sum_{E_2} X_i/n_2$, then mean value m of Y_n equals to m_1-m_2 and variance of Y_n σ_n^2 equals to $\sigma_1^2/n_1+\sigma_2^2/n_2$. Under the procedure $\mathcal{P}$ we have shown in lemma 3 that $\min(n_1, n_2)\to\infty$ as $n\to\infty$, therefore our σ_n^2 converges to zero.

Now we assume $m=m_1-m_2=0$, then Y_n/σ_n is a ramdom variable of normal population $N(0,1)$, so that Y_n^2/σ_n^2 is a random variable of χ^2 population with one degree of freedom. We put it

$$\frac{Y_n^2}{\sigma_n^2}=\chi_1^2. \tag{4.12}$$

Therefore $Y_n^2=\sigma_n^2\chi_1^2$ then Y_n^2 is distributed as χ^2 distribution with constant coefficient σ_n^2, the mean value of Y_n^2 equals to σ_n^2 and the variance equals to $2\sigma_n^4$.

In the following lines we assume $m=m_1-m_2\neq 0$, then by the equality

$$\frac{Y_n^2-m^2}{2m\sigma_n}-\frac{Y_n-m}{\sigma_n}=\frac{(Y_n-m)^2}{2m\sigma_n} \tag{4.13}$$

we can get the asymptotic behavior of Y_n^2. As $(Y_n-m)^2$ is asymptotically equal to zero in higher order as compared with (Y_n-m), $Y_n^2-m^2/2m\sigma_n$ is asymptotically equal to $(Y_n-m)/\sigma_n$ in probability, where the random variable $(Y_n-m)/\sigma_n$ is a random variable of normal population $N(0,1)$. In the following lines we put the random variable as Z, then we have

$$\frac{Y_n^2-m^2}{2m\sigma_n}\sim Z \tag{4.14}$$

in probability for sufficiently large n. Therefore, for sufficiently large n, Y_n^2 is asymptotically distributed as $2m\sigma_nZ+m^2$ in probability, then this is a random variable of normal population $N(m^2, 4m^2\sigma_n^2)$. Hence under the procedure $\mathcal{P}$ we have the following

LEMMA 1. *Under the procedure $\mathcal{P}$, if $m_1\neq m_2$, then $(\sum_{E_1} X_i/n_1-\sum_{E_2} X_i/n_2)^2$ is asymptotically normally distributed with mean $(m_1-m_2)^2$ and variance $4(m_1-m_2)^2(\sigma_1^2/n_1+\sigma_2^2/n_2)$ in probability. And if $m_1=m_2$, then the random variable is exactly distributed as χ^2 distribution with one degree of freedom having a coefficient $\sigma_1^2/n_1+\sigma_2^2/n_2$.*

Note that the order of the variance of the random variable $(\sum_{E_1} X_i/n_1 - \sum_{E_2} X_i/n_2)^2$ in the case $m_1 \neq m_2$ is given by

$$4(m_1-m_2)^2\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)=O\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)=O\left(\frac{\sigma_1^2}{n\lambda}+\frac{\sigma_2^2}{n(1-\lambda)}\right)=O\left(\frac{1}{n}\right). \tag{4. 15}$$

And, in the case $m_1=m_2$, we have the variance of the random variable is given by (4. 12)

$$\begin{aligned}\operatorname{Var}\left\{\left(\frac{\sum_{E_1} X_i}{n_1}-\frac{\sum_{E_2} X_i}{n_2}\right)^2\right\}&=2\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)^2\\ &=O\left(\left\{\frac{\sigma_1^2}{n\lambda}+\frac{\sigma_2^2}{n(1-\lambda)}\right\}^2\right)=O\left(\frac{1}{n^2}\right):\end{aligned} \tag{4. 16}$$

By the equation (4. 11) our function $S_n(\hat{\theta}_n, \tilde{\theta}_n)/n$ is given as follows.

$$\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{n}=\frac{1}{2}\frac{(n_1/n)(1-n_1/n)}{(n_1/n)(\sigma_2^2-\sigma_1^2)+\sigma_1^2}\left(\frac{\sum_{E_1} X_i}{n_1}-\frac{\sum_{E_2} X_i}{n_2}\right)^2. \tag{4. 17}$$

In the right hand of this equation only $(\sum_{E_1} X_i/n_1-\sum_{E_2} X_i/n_2)^2$ is a random variable. Using the result of lemma 1 we can get the asymptotic behavior of $S_n(\hat{\theta}_n, \tilde{\theta}_n)/n$ as follows.

In the case $m_1 \neq m_2$, $S_n(\hat{\theta}_n, \tilde{\theta}_n)/n$ is asymptotically normally distributed with mean

$$\frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)(m_1-m_2)^2 \tag{4. 18}$$

and variance

$$\begin{aligned}&\left(\frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)\right)^2 4(m_1-m_2)^2\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)\\ &=\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)^2\frac{n_2\sigma_1^2+n_1\sigma_2^2}{n_1n_2}(m_1-m_2)^2\\ &=\frac{1}{n^2}\frac{n_1n_2}{n_1\sigma_2^2+n_2\sigma_1^2}(m_1-m_2)^2\\ &=\frac{1}{n}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)(m_1-m_2)\\ &=O\left(\left\{\frac{1}{n}\frac{\lambda(1-\lambda)}{\lambda\sigma_2^2+(1-\lambda)\sigma_1^2}(m_1-m_2)^2\right\}\right)=O\left(\frac{1}{n}\right).\end{aligned} \tag{4. 19}$$

And if $m_1=m_2$, then by (4. 12) our $S_n(\hat{\theta}_n, \tilde{\theta}_n)/n$ becomes

$$\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{n} = \frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2$$

$$(4.20)\qquad = \frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)\chi_1^2$$

$$= \frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)\frac{1}{n}\left(\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\Big/\frac{n_1}{n}\frac{n_2}{n}\right)\chi_1^2$$

$$= \frac{1}{2n}\chi_1^2.$$

Therefore if $m_1=m_2$, our $S_n(\hat{\theta}_n, \tilde{\theta}_n)/n$ is distributed as χ^2 distribution with one degree of freedom having a coefficient number $1/2n$ under the procedure $\mathcal{P}$. Hence we have the next.

THEOREM. *Under the procedure $\mathcal{P}$ if $m_1 \neq m_2$, then $S_n(\hat{\theta}_n, \tilde{\theta}_n)/n$ is asymptotically normally distributed with mean*

$$\frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)(m_1-m_2)^2$$

and variance

$$\frac{1}{n}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)(m_1-m_2)^2.$$

And if $m_1=m_2$, then

$$\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{n} = \frac{1}{2n}\chi_1^2$$

is distributed exactly χ^2 distribution with one degree of freedom having a coefficient number $1/2n$.

Finally we shall show the limit value of $S_n(\hat{\theta}_n, \tilde{\theta}_n)/n$ under the procedure $\mathcal{P}$. By lemma 3 we have seen the fact $n_1/n \to \lambda$ under the procedure $\mathcal{P}$. And by the fact $0<\lambda<1$ we have $\min(n_1, n_2) \to \infty$ as $n\to\infty$ under the procedure $\mathcal{P}$. Therefore by lemma 1 we have $\hat{\theta}_n \to \theta$ as $n\to\infty$ with probability one under the procedure $\mathcal{P}$. Hence, under the procedure $\mathcal{P}$,

$$(4.21)\qquad \frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{n} = \frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2^2+\frac{n_2}{n}\sigma_1^2\right)\right)\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2$$

$$\to \frac{1}{2}\frac{\lambda(1-\lambda)}{\lambda\sigma_2^2+(1-\lambda)\sigma_1^2}(m_1-m_2)^2$$

as $n\to\infty$ with probability one.

4. 4. In this place we consider the limiting property and the asymptotic behavior of the logarithm of the likelihood ratio per unit cost with respect to the hypothesis $m_1=m_2$: $S_n(\hat{\theta}_n, \tilde{\theta}_n)/\sum_{i=1}^n C^{(i)}$.

First we shall show the limiting property of the function $S_n(\hat{\theta}_n, \tilde{\theta}_n)/\sum_{i=1}^n C^{(i)}$. Under the procedure $\mathcal{P}$ we have

$$\lim_{n\to\infty}\frac{n_1}{n}=\lambda, \qquad P\left\{\lim_{n\to\infty}\frac{\sum_{E_1}X_i}{n_1}=m_1\right\}=1 \quad \text{and} \quad P\left\{\lim_{n\to\infty}\frac{\sum_{E_2}X_i}{n_2}=m_2\right\}=1$$

as we have seen in preceding 4. 3. Our function can be written in the form

$$\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=\frac{n}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}\sigma_2{}^2+\frac{n_2}{n}\sigma_1{}^2\right)\right)\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2\Big/(n_1C_1+n_2C_2)$$

$$=\frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}C_1+\frac{n_2}{n}C_2\right)\left(\frac{n_1}{n}\sigma_2{}^2+\frac{n_2}{n}\sigma_1{}^2\right)\right)\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2.$$

Then under the procedure $\mathcal{P}$ we can easily get

$$\lim_{n\to\infty}\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=\frac{1}{2}\frac{\lambda(1-\lambda)(m_1-m_2)^2}{(\lambda C_1+(1-\lambda)C_2)(\lambda\sigma_2{}^2+(1-\lambda)\sigma_1{}^2)} \tag{4. 23}$$

with probability one.

In the following we shall show the asympiotic behavior of our function $S_n(\hat{\theta}_n, \tilde{\theta}_n)/\sum_{i=1}^n C^{(i)}$. Under the procedure $\mathcal{P}$, by lemma 1 in preceding 4. 3, if $m_1\neq m_2$ then $(\sum_{E_1}X_i/n_1-\sum_{E_2}X_i/n_2)^2$ asymptotically distributed normally with mean $(m_1-m_2)^2$ and variance $4(m_1-m_2)^2(\sigma_1{}^2/n_1+\sigma_2{}^2/n_2)$ in probability and if $m_1=m_2$ then we have given $(\sum_{E_1}X_i/n_1-\sum_{E_2}X_i/n_2)^2=(\sigma_1{}^2/n_1+\sigma_2{}^2/n_2)\chi_1{}^2$.

In the equality (4. 21) the coefficient of $(\sum_{E_1}X_i/n_1-\sum_{E_2}X_i/n_2)^2$ is a function of n because of the properties (3. 19) and (3. 20) of the procedure $\mathcal{P}$. Therefore if $m_1\neq m_2$ the function $S_n(\hat{\theta}_n, \tilde{\theta}_n)/\sum_{i=1}^n C^{(i)}$ is asymptotically distributed normally with mean

$$\frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}C_1+\frac{n_2}{n}C_2\right)\left(\frac{n_1}{n}\sigma_2{}^2+\frac{n_2}{n}\sigma_2{}^2\right)\right)(m_1-m_2)^2$$

and variance

$$\left\{\frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}C_1+\frac{n_2}{n}C_2\right)\left(\frac{n_1}{n}\sigma_2{}^2+\frac{n_2}{n}\sigma_1{}^2\right)\right)\right\}^2 4(m_1-m_2)^2\left(\frac{\sigma_1{}^2}{\sigma_1}+\frac{\sigma_2{}^2}{n_2}\right)$$

$$=\frac{1}{n}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}C_1+\frac{n_2}{n}C_2\right)^2\left(\frac{n_1}{n}\sigma_2{}^2+\frac{n_2}{n}\sigma_1{}^2\right)\right)(m_1-m_2)^2 \tag{4. 24}$$

$$\sim\frac{1}{n}\frac{\lambda(1-\lambda)}{(\lambda C_1+(1-\lambda)C_2)^2(\lambda\sigma_2{}^2+(1-\lambda)\sigma_1{}^2)}(m_1-m_2)^2=O\left(\frac{1}{n}\right).$$

And if $m_1=m_2$ then $S_n(\hat{\theta}_n, \tilde{\theta}_n)/\sum_{i=1}^n C^{(i)}$ can be written as following

$$\frac{S_n(\hat{\theta}_n, \tilde{\theta}_n)}{\sum_{i=1}^n C^{(i)}}=\frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}C_1+\frac{n_2}{n}C_2\right)\left(\frac{n_1}{n}\sigma_2{}^2+\frac{n_2}{n}\sigma_1{}^2\right)\right)\left(\frac{\sum_{E_1}X_i}{n_1}-\frac{\sum_{E_2}X_i}{n_2}\right)^2$$

$$=\frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}C_2+\frac{n_2}{n}C_2\right)\left(\frac{n_1}{n}\sigma_2{}^2+\frac{n_2}{n}\sigma_1{}^2\right)\right)\left(\frac{\sigma_1{}^2}{\sigma_1}+\frac{\sigma_2{}^2}{n_2}\right)\chi_1{}^2$$

$$=\frac{1}{2n}\left(1\Big/\left(\frac{n_1}{n}C_1+\frac{n_2}{n}C_2\right)\right)\chi_1{}^2.$$

Therefore if $m_1=m_2$ then $S_n(\hat{\theta}_n, \tilde{\theta}_n)/\sum_{i=1}^n C^{(i)}$ is exactly distributed as χ^2 distribution with one degree of freedom which has a coefficient $(1/2n)(1/(n_1C_1/n+n_2C_2/n))$. Therefore we have the next theorem as to be proved.

THEOREM. *Under the procedure* $\mathcal{P}$ *if* $m_1\neq m_2$, *then* $S_n(\hat{\theta}_n, \tilde{\theta}_n)/\sum_{i=1}^n C^{(i)}$ *is asymptotically normally distributed with mean*

$$\frac{1}{2}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}C_1+\frac{n_2}{n}C_2\right)\left(\frac{n_1}{n}\sigma_2{}^2+\frac{n_2}{n}\sigma_1{}^2\right)\right)(m_1-m_2)^2$$

and variance

$$\frac{1}{n}\left(\frac{n_1}{n}\frac{n_2}{n}\Big/\left(\frac{n_1}{n}C_1+\frac{n_2}{n}C_2\right)^2\left(\frac{n_1}{n}\sigma_2{}^2+\frac{n_2}{n}\sigma_1{}^2\right)\right)(m_1-m_2)^2.$$

And if $m_1=m_2$, *then exactly distributed as* χ^2 *distribution with one degree of freedom and the ratio can be represented by*

$$\frac{1}{2n}\left(1\Big/\frac{n_1}{n}C_1+\frac{n_2}{n}C_2\right)\Big)\chi_1{}^2,$$

where $\chi_1{}^2$ *is a* χ^2 *distributed random variable with one degree of freedom.*

ACKNOWLEDGEMENT. I express my gratitude to Prof. Dr. K. Kunisawa for his useful advice.

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] KAWAMURA, K., Asymptotic behavior of sequential design with costs of experi-

ments (The case of normal distribution). Kōdai Math. Sem. Rep. **17** (1965), 48–52.
[4] KAWAMURA, K., Asymptotically most informative procedure in the case of exponential families. Kōdai Math. Sem. Rep. **19** (1967), 61–74.
[5] KULLBACK, S., Information theory and statistics. Wiley (1959).

DEPARTMENT OF MATHEMATICS,
TOKYO INSTITUTE OF TECHNOLOGY.