SUCCESIVE MULTIPLE DECISION PROCEDURES FOR ORDERED PARAMETERS

Tasaka, Masao
Faculty of Engineering, Osaka University

https://doi.org/10.5109/13068

出版情報：統計数理研究. 15 (3/4), pp.1-20, 1973-03. Research Association of Statistical Sciences
バージョン：
権利関係：

SUCCESSIVE MULTIPLE DECISION PROCEDURES FOR ORDERED PARAMETERS

By

Masao TASAKA*

(Received November 25, 1971)

1. Statement of problem.

We consider a multiple decision problem for deciding an ordering of k unknown parameters $\theta_1, \theta_2, \cdots, \theta_k$. Let $(i_1, i_2, \cdots, i_k)$ be one of the $k!$ permutations of $(1, 2, \cdots, k)$. Suppose that we are concerned with a set of $k!\sum_{j=1}^{k}(j!)^{-1}$ hypotheses which are defined by

$$\left.\begin{aligned} &H_0:\ \theta_1=\theta_2=\cdots=\theta_k \\ &H_{i_1\cdots i_r0}:\ \theta_{i_1}\geqq\cdots\geqq\theta_{i_r}\geqq\theta_{i_{r+1}}=\cdots=\theta_{i_k} \qquad (r=1,2,\cdots,k-1) \end{aligned}\right\} \tag{1-1}$$

where for at last one i_s $(s=1, \cdots, r)$, the strict inequality $\theta_{i_s}>\theta_{i_{s+1}}$ holds in $H_{i_1\cdots i_r0}$. Let $D.$ be the decision to accept the corresponding hypothesis $H.$, $P_r\{D.|H'.\}$ be the probability of making decision $D.$ when $H'.$ is true, and $P_r\{D.|D'.|H''\}$ be the conditional probability of making decision $D.$, given the decision $D'.$ when $H''.$ is true. Under the situation that we can be sure of the existence of the true hypothesis among the above set of $k!\sum_{j=1}^{k}(j!)^{-1}$ hypotheses (1-1), we naturally wish to find a multiple decision procedure which maximizes the probabilities of correct decisions, $Pr\{D_{i_1\cdots i_r0}|H_{i_1\cdots i_r0}\}$, simulataneously. This problem is too vague to be handled as the amount of knowledge on the ture situation of ordered parameters is small, but we have to start somewhere; in fact, further specialization of the models and/or decision procedure itself is usually necessary.

We now consider the following type of hypotheses.

$$H_{i_1\cdots i_r}:\ \theta_{i_1}\geqq\cdots\geqq\theta_{i_r}\geqq\theta_{i_{r+1}}, \cdots, \theta_{i_k}, \qquad (r=1,2,\cdots,k-1). \tag{1-2}$$

For two sets (1-1) and (1-2) of hypotheses, we have the following logical relations. If any hypothesis $H_{i_1\cdots i_r}(H_{i_1\cdots i_r0})$ for fixed $i_1, \cdots, i_k$ is true, the preceding hypotheses $H_{i_1\cdots i_{r-1}}, \cdots, H_{i_1}(H_{i_1\cdots i_{r+1}0}, \cdots, H_{i_1\cdots i_{k-1}0})$ must be true, and if any hypothesis is false the succeeding ones must be false, that is,

$$\left.\begin{aligned} &H_{i_1}\supseteqq H_{i_1i_2}\supseteqq\cdots\supseteqq H_{i_1\cdots i_r} \\ \text{and}\quad &H_{i_1\cdots i_r0}\subseteqq H_{i_1\cdots i_{r+1}0}\subseteqq\cdots\subseteqq H_{i_1\cdots i_{k-1}0}. \end{aligned}\right\} \tag{1-3}$$

* Faculty of Engineering, Osaka University, Osaka.

And the two sets of hypotheses (1-1) and (1-2) are related by

$$H_{i_1\cdots i_r} \supseteqq H_{i_1\cdots i_r0}\,. \tag{1-4}$$

We have

$$Pr\{D_{i_1\cdots i_r0}\} = Pr\{D_{i_1}\}Pr\{D_{i_1i_2}|D_{i_1}\} \cdots Pr\{D_{i_1\cdots i_r}|D_{i_1\cdots i_{r-1}}\}Pr\{D_{i_1\cdots i_r0}|D_{i_1\cdots i_r}\}\,. \tag{1-5}$$

Further we consider the following hypotheses

$$\left.\begin{array}{l} H_{(i_1\cdots i_r)0}:\ \theta_{i_{r+1}} = \cdots = \theta_{i_k} \\ H_{(i_1\cdots i_r)i_{r+1}\cdots i_s0}:\ \theta_{i_{r+1}} \geqq \cdots \geqq \theta_{i_s} \geqq \theta_{i_{s+1}} = \cdots = \theta_{i_k} \\ H_{(i_1\cdots i_r)i_{r+1}\cdots i_s}:\ \theta_{i_{r+1}} \geqq \cdots \geqq \theta_{i_s} \geqq \theta_{i_{s+1}}, \cdots, \theta_{i_k}\,. \end{array}\right\} \tag{1-6}$$

Now if the condition

(C1-1) $$Pr\{D_{i_1\cdots i_r}|D_{i_1\cdots i_{r-1}}\} = Pr\{D_{(i_1\cdots i_{r-1})i_r}\}$$

holds, then we obtain

$$Pr\{D_{i_1\cdots i_r0}\} = Pr\{D_{i_1}\}Pr\{D_{(i_1)i_2}\} \cdots Pr\{D_{(i_1\cdots i_{r-1})i_r}\}Pr\{D_{(i_1\cdots i_r)0}\}\,. \tag{1-7}$$

We shall desire to find a successive multiple decision procedure which maximizes the probability of correct decision. In case when, for instance, $H_{j_1\cdots j_s0}$ is true, our requirement is, in terms of our notation,

(C1-2) $$\underset{(i_1\cdots i_r)}{\text{Max}}\, Pr\{D_{i_1\cdots i_r}|H_{j_1\cdots j_r\cdots j_s0}\} = Pr\{D_{j_1\cdots j_r}|H_{j_1\cdots j_r\cdots j_s0}\}$$
$$(r=1, 2, \cdots, s\,;\ 1\leqq s\leqq k-1)$$

where the maximum is the one for all possible combinations of $(i_1, \cdots, i_r)$.

Let $\Phi.$ be a class of possible decision functions $\varphi.$ satisfying the conditions (C1-1) and (C1-2). It follows from (1-7) that

$$\begin{aligned} &\underset{\Phi_{i_1\cdots i_s0}}{\text{Max}}\ \log Pr\{D_{i_1\cdots i_s0}|H_{12\cdots s0}\} \\ &\quad = \underset{\Phi_{i_1\cdots i_s0}}{\text{Max}}\,[\log Pr\{D_{i_1}|H_{12\cdots s0}\} + \log Pr\{D_{(i_1)i_2\cdots i_s0}|H_{12\cdots s0}\}] \\ &\quad = \underset{\Phi_{i_1}}{\text{Max}}\,[\log Pr\{D_{i_1}|H_{12\cdots s0}\} + \underset{\Phi_{i_2\cdots i_s0}}{\text{Max}}\ \log Pr\{D_{(i_1)i_2\cdots i_s0}|H_{12\cdots s0}\}] \\ &\quad = [\log Pr\{D_1|H_{12\cdots s0}\} + \underset{\Phi_{i_2\cdots i_s0}}{\text{Max}}\ \log Pr\{D_{(1)i_2\cdots i_s0}|H_{12\cdots s0}\}]\,. \end{aligned} \tag{1-8}$$

However, it is difficult in general to satisfy the requirement (C1-2), because it would be expected to satisfy the relation

$$\underset{\Phi_{i_2\cdots i_s0}}{\text{Max}}\ Pr\{D_{(i)i_2\cdots i_s0}|H_{12\cdots s0}\} \geqq \underset{\Phi_{i_2\cdots i_s0}}{\text{Max}}\ Pr\{D_{(1)i_2\cdots i_s0}|H_{12\cdots s0}\} \quad (\text{for } i\in(2, 3, \cdots, s))\,. \tag{1-9}$$

In such a situation, it could not be desired to find a successive multiple decision procedure which maximizes the probability of correct decision. Therefore, instead of (C1-2) we have to be satisfied with the requirement under any fixed $(i_1, \cdots, i_r)$:

(C1-3) $$\underset{i_s\in(i_{r+1},\cdots,i_k)}{\text{Max}}\ Pr\{D_{(i_1\cdots i_r)i_s}|H_{j_1\cdots j_r\cdots j_s0}\} = Pr\{D_{(i_1\cdots i_r)\rho}|H_{j_1\cdots j_r\cdots j_s0}\}$$

where

$$\rho = \underset{r \text{ in } j_r}{\text{Max}} \{(j_1, \cdots, j_r, \cdots, j_s) - (i_1, \cdots, i_r) \cap (j_1, \cdots, j_r, \cdots, j_s)\} .$$

It would be expected that there are a number of possible decision functions satisfying the condition (C1-3).

The formulation of our problem can be modified to include the case where the standard value or a control population is present. In this case the hypotheses (1-1) should be change to

$$\begin{aligned} &H_0: \theta_1 = \theta_2 = \cdots = \theta_k = \theta_0 \\ &H_{i_1 \cdots i_r 0}: \theta_{i_1} \geqq \cdots \geqq \theta_{i_r} \geqq \theta_{i_{r+1}} = \cdots = \theta_{i_k} = \theta_0 \qquad (r = 1, 2, \cdots, k-1) \\ &H_{i_1 \cdots i_k}: \theta_{i_1} \geqq \theta_{i_2} \geqq \cdots \geqq \theta_{i_k} > \theta_0 \end{aligned} \tag{1-10}$$

and the number of hypotheses is $k!\left(1 + \sum_{j=1}^{k} (j!)^{-1}\right)$. The discussion made above is also valid somewhat in this modified formulation.

In Section 2 and 3, we propose a successive multiple decision procedure which satisfies the condition (C1-1) and would have reasonable power in the likelihood ratio test criterion. In Section 4 and 5 we consider linear hypotheses model. A slippage problem in the general linear hypotheses model is discussed in Section 4 for specifying the operating characteristics of a successive multiple decision procedure which is introduced in Section 5. An example, in Section 6, will show that there does not exists the successive decision procedure which is optimum in the sense that uniformly maximizes the probability of correct decision.

2. Nested induced test.

In the problem of testing a given hypothesis about an unknown vector parameter, a method of testing has suggested by Hogg [6], and Daroch and Silvey [3] indicated that this method, which, according to Seber [11], we shall call the nested induced test, will have reasonable power if the likelihood ratio statistic is used as the test criterion. If $H_{12 \cdots s0}$ is a prior given hypotheses, then it is possible to use the nested induced test for testing $H_{12 \cdots s0}$. However, since we are treating our decision problem under the incomplete information that there exists a true hypothesis among the set of $k! \sum_{j=1}^{k} (j!)^{-1}$ hypotheses (1-1), we cannot directly use the nested induced test itself, but a successive method for multiple decision may be suggested by examining some properties of nested induced test as shown in this section.

Suppose for each $i = 1, 2, \cdots, k$ the random variable X_i has the density function $f(x; \theta_i)$, and random samples $x_1, \cdots, x_k$ are independently taken from the distribution of $X_1, \cdots, X_k$. Let I be an interval of real numbers, $\Omega = \{\theta_1, \cdots, \theta_k; \theta_i \in I (i = 1, \cdots, k)\}$ be k dimensional Euclidean space R^k or a subspace of R^k, $\omega.$ denotes the subspace of Ω that the hypothesis $H.$ is satisfied, and G be a general stochastic model in which a vector parameter, $\theta = (\theta_1, \cdots, \theta_k)$ is known to belong to Ω, that is, $\theta \in \Omega$. Let $L[\theta]$ be the likelihood function

$$L[\theta]=\prod_{j=1}^{k} f(x_j:\theta) \tag{2-1}$$

and difine the likelihood ratio by

$$L[H..|H.]=\sup_{\theta\in\omega.\cap\omega..} L[\theta]/\sup_{\theta\in\omega.} L[\theta]\,, \tag{2-2}$$

then we have

$$L[H_{i_1\cdots i_r0}|G]=L[H_{i_1}|G]L[H_{i_1i_2}|H_{i_1}]\cdots L[H_{i_1\cdots i_r}|H_{i_1\cdots i_{r-1}}]L[H_{i_1\cdots i_r0}|H_{i_1\cdots i_r}]\,. \tag{2-3}$$

As pointed out in Seber [11] and Darroch and Silvey [3], if each of the likelihood ratios on the right hand side is near one, then left hand side will be near one also. Therefore, as the likelihood ratio test has, in general, reasonable power, then, the following nested induced test will also have reasonable power: we accept $H_{i_1\cdots i_r0}$ against G only if all of the hypotheses in the sequence of the tests

(i) H_{i_1} against G,

(ii) $H_{i_1i_2}$ against H_{i_1},

....

(r) $H_{i_1\cdots i_r}$ against $H_{i_1\cdots i_{r-1}}$,

($r+1$) $H_{i_1\cdots i_r0}$ against $H_{i_1\cdots i_r}$

are accepted.

We shall now examine some properties of the nested induced test mentioned above. In order to do so we shall now specify the nature of the density function $f(x;\theta)$:

(C2-1) $f(x;\theta)$ has monotone likelihood ratio in $T(x)$, that is, there exists a real-valued function $T(x)$ such that for any $\theta<\theta'$ the distribution $F(x;\theta)$ and $F(x;\theta')$ are distinct, and the ratio $f(x;\theta')/f(x;\theta)$ is a nondecreasing function of $T(x)$.

We have the following.

LEMMA 2-1. *If there exists a unique θ which maximizes $L[\theta]$ in the domain Ω (and ω_i), then under the condition* (C2-1), *the likelihood ratio $L[H_i|G]$ is a nondecreasing function of $T(x_i)$ and for each $j=1,2,\cdots,i-1,i+1,\cdots,k$ a nonincreasing function of $T(x_j)$.*

PROOF. Let $\hat\theta=(\hat\theta_1,\cdots,\hat\theta_k)$ be a maximum likelihood estimate under H_i, then it maximizes $L[\theta]$ in the domain ω_i. Then it should be holds that

$$\hat\theta_i>\mathrm{Max}\,(\hat\theta_1,\cdots,\hat\theta_{i-1},\hat\theta_{i+1},\cdots,\hat\theta_k)\,. \tag{2-4}$$

Let $\hat\theta'=(\hat\theta'_1,\cdots,\hat\theta'_k)$ be a maximum likelihood estimate of which maximizes $L[\theta]$ in the domain Ω. Assume now we have $\hat\theta'_i>\hat\theta_i$. Then it follows from (2-4) that

$$(\hat\theta_1,\cdots,\hat\theta_{i-1},\hat\theta'_i,\hat\theta_{i+1},\cdots,\hat\theta_k)\in\omega_i\,. \tag{2-5}$$

On the other hand, from the uniqueness of maximum likelihood estimate we have

$$f(x_i;\hat\theta'_i)>f(x_i;\hat\theta_i)$$

and thus

$$f(x_i\,;\,\hat{\theta}'_i)\prod_{\substack{j=1\\ j\neq i}}^{k} f(x_j\,;\,\hat{\theta}_j) > \prod_{j=1}^{k} f(x_j\,;\,\hat{\theta}_j)\,.$$

From (2-5) this is a contradiction, and therefore $\hat{\theta}'_i \leqq \hat{\theta}_i$ has to be true, as was to be proved. Similarly, it can be shown that $\hat{\theta}'_j \geqq \hat{\theta}_j$ $(j=1,2,\cdots,i-1,i+1,\cdots,k)$. This completes the proof.

For indicating clearly the existence and representation of maximum likelihood estimate of θ it is necessary to restrict the class of distribution. It is contained in the works of Brunk [1], van Eeden [13] and others who investigated the problem for estimating monotone parameters. We are now going to study under the formulation which is similar to the one described by Robertson and Waltman [10]. Suppose that the real parameter family of densities $f(x\,;\,\theta)$ has the following properties:

(C2-2) $f(x\,;\,\theta)$ has a support S which is the same for all $\theta \in I$,

(C2-3) for each x in S, $f(x\,;\,\theta)$ is a continuous function of θ on I,

(C2-4) $f(x\,;\,\theta)$ is a strictly unimodal function of θ in I,

that is, there exists a number M in I which is a unique mode of $f(x\,;\,\theta)$.

In Brunk [1], Robertson and Waltman [10] it is shown that under the conditions (C2-2), (C2-3) and (C2-4) there is a maximum likelihood estimate of θ in $\omega_{12\cdots k}$ and it may be obtained as follows:

Let M_i be the mode of $f(x_i\,;\,\theta)$. If $M_1 \geqq M_2 \geqq \cdots \geqq M_k$, then $\hat{\theta}_i = M_i$, $i=1,2,\cdots,k$. On the other hand, if for some i we have $M_i < M_{i+1}$, then the i-th and $i+1$-th samples are pooled and M_{ii+1} is the mode of $f(x_i\,;\,\theta)\cdot f(x_{i+1}\,;\,\theta)$. Furthermore $M_{ii+1} = M_j$ for $j<i$ and $M_{ii+1} = M_{j+1}$ for $j>i$. This procedure is repeated until a set of monotone nonincreasing set of modes, M, is obtained. They showed that $\sup_{\theta \in \omega i} L[\theta]$ is obtained as $L[M]$ from the sample by this procedure, which we shall call $R-M$ procedure.

LEMMA 2-2. *Under the conditions* (C2-2), (C2-3) *and* (C2-4), *the likelihood ratio* $L[H_i|G]$ *is a nondecreasing function of* M_i *and a nonincreasing function of* $M_1,\cdots,M_{i-1}$, $M_{i+1},\cdots,M_k$.

PROOF. For the sake of simplicity and without loss of generality we consider the case of $i=1$ and suppose that a maximum likelihood estimate of θ in Ω is given by $\hat{\theta}'_j = M_j$ $(j=1,2,\cdots,k)$ from the sample $(x_1,\cdots,x_k)$ and

$$M_s \geqq M_{s-1} \geqq \cdots \geqq M_2 \geqq M_1 \geqq M_{s+1} \geqq \cdots \geqq M_k$$

is obtained. According to $R-W$ procedure it is easily shown that a maximum likelihood estimate of θ in ω_i is given by

$$\hat{\theta}_j = \begin{cases} M_{12\cdots s} & \text{for } j=1,2,\cdots,s \\ M_j & \text{for } j=s+1,\cdots,k \end{cases}$$

and thus

$$L[H_1|G] = \prod_{j=1}^{s} f(x_j\,;\,M_{12\cdots s})/f(x_j\,;\,M_j) \tag{2-6}$$

holds. On the other hand, it follows from the condition (C2-4) that M_{ii+1} is between M_i and M_{i+1}, and thus

$$M_s \geqq M_{12\cdots s} \geqq M_1 . \qquad (2\text{-}7)$$

It is proved from (2-6) and (2-7) that $L[H_1|G]$ is a nondecreasing function of M_1. It follows from the condition (C2-4) that

$$M_2 \geqq M_{12} \geqq M_1 ,$$

and hence M_{12} does not decrease as M_2 increases. Similarly

$$M_3 \geqq M_{123} \geqq M_{12} ,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$M_s \geqq M_{12\cdots s} \geqq M_{12\cdots s-1} ,$$

and hence $M_{12\cdots s}$ does not decrease as $M_2, \cdots, M_{s-1}$ and M_s increase respectively. On the other hand, since

$$M_{2\cdots s} \geqq M_{12\cdots s} \geqq M_1$$

holds, $L[H_1|G]$ is a nonincreasing function of $M_{2\cdots s}$. In this maner we can prove the argument.

An important class of families of distributions that satisfy the condition (C2-1) is the one-parameter exponential families

$$f(x;\theta) = C(\theta)e^{Q(\theta)T(x)}h(x)$$

where Q is strictly monotone. Brunk [1] considered the present problem when $f(x;\theta)$ belongs to a certain exponential family of distributions. It is known from Lemma 2-1 and 2-2 that if $T(x_i) = M_i$ $(i = 1, 2, \cdots, k)$, then the condition (C2-1) and the assumption of uniqueness of maximum likelihood estimate in ω_i are equivalient to the condition (C2-4).

THEOREM 2-1. *Under the conditions* (C2-2), (C2-3) *and* (C2-4) *the likelihood ratio* $L[H_{i_1\cdots i_{r-1}i_r}|H_{i_1\cdots i_{r-1}}]$ *is independent of* $x_{i_1},\cdots,x_{i_{r-1}}$, *a nondecreasing function of* M_{i_r} *and for each* $j = r+1, \cdots, k$ *a nonincreasing function of* M_j.

PROOF. Consider the following hypotheses

$$H_{(i_1\cdots i_{r-1})i_r} : \theta_{i_r} \geqq \theta_{i_{r+1}}, \cdots, \theta_{i_k} \qquad (r = 2, 3, \cdots, k-1) . \qquad (2\text{-}8)$$

It is easily shown from the procedure for obtaining a maximum likelihood estimate of θ in $\omega_{i_1\cdots i_k}$ that

$$L[H_{i_1\cdots i_{r-1}i_r}|H_{i_1\cdots i_{r-1}}] = L[H_{i_1\cdots i_{r-1}i_r}|\bar{G}_{k-r+1}]$$

where $\bar{G}_{k-r+1}$ is a general stochastic model in which a vector parameter $\bar{\theta} = (\theta_{i_r}, \theta_{i_{r+1}}, \cdots, \theta_{i_k})$ is known to belong $\{\bar{\theta}\,;\,\theta_{i_j} \in I, j = r, r+1, \cdots, k\}$. By using Lemma 2-2, this completes the argument.

(2-3) and Theorem 2-1 indicate that the nested induced test for a given hypothesis $H_{i_1\cdots i_r0}$ satisfies the condition (C1-1), and is a sequence of likelihood ratio test based on likelihood ratio which has the property of Lemma 2-2, that is,

$$L[H_{i_1\cdots i_r0}|G] = L[H_{i_1}|G]L[H_{(i_1)i_2}] \cdots L[H_{(i_1\cdots i_{r-1})i_r}]L[H_{(i_1\cdots i_r)0}] \qquad (2\text{-}9)$$

3. Successive multiple decision procedure.

In our problem the true hypothesis is one of $k!\sum_{j=1}^{k}(j!)^{-1}$ hypotheses (1-1) and no a priori knowledge available to us. So we wish to propose a successive test procedure for giving mutiple decision in this problem by using the results of the preceding Section 2. In order to do this, we first examine the properties of the following likelihood ratio test:

if
$$L[H_{i_1}|G] \geqq \lambda_1 \tag{3-1}$$
holds, then we select a decision D_{i_1}, where λ_1 is a value determined by
$$Pr\{L[H_{i_1}|G] < \lambda_1\} < \alpha_1 \qquad (i_1 = 1, 2, \cdots, k) \tag{3-2}$$
and α_1 is a preassigned value with the condition $\alpha_1 \in (0, 1)$. It is clear that if
$$M_{i_1} \geqq M_{i_2} \geqq \cdots \geqq M_{i_k} \tag{3-3}$$
is given from the sample $(x_1, \cdots, x_k)$, and a decision D_{i_r} is selected for an integer r, then the decisions $D_{i_1}, \cdots, D_{i_{r-1}}$ should be necessarily selected.

Now we define
$$H_{\overline{i_1 \cdots i_r} \cdots \overline{i_s \cdots i_t}} : \theta_{i_1} = \cdots = \theta_{i_r} > \cdots > \theta_{i_s} = \cdots = \theta_{i_t} > \theta_{i_{t+1}}, \cdots, \theta_{i_k}\,. \tag{3-4}$$
Then we have
$$H_{i_1} \cap \cdots \cap H_{i_r} = H_{\overline{i_1 \cdots i_r}}\,. \tag{3-5}$$
Hence we know that if $D_{i_1}, \cdots, D_{i_r}$ are jointly selected (and $D_{i_{r+1}}, \cdots, D_{i_k}$ are not selected) according to the above likelihood ratio test (3-1), then we select the decision $D_{\overline{i_1 \cdots i_r}}$.

By considering of this and (2-9), we shall propose the following successive multiple decision procedure:

Let λ_r be a value satisfied by
$$Pr\{L[H_{i_1 \cdots i_{r-1} i_r} | H_{i_1 \cdots i_{r-1}}] < \lambda_r\} \leqq \alpha_r\,, \tag{3-6}$$
and α_r be a preassigning value with the condition $\alpha_r \in (0, 1)$. When the ordering modes of $f(x_i\,;\theta_i)$ $(i=1, \cdots, k)$ according to (3-3) are observed from the sample $(x_1, \cdots, x_k)$,

(1.) if $L[H_{i_j}|G] \geqq \lambda_1 \qquad (j=1, \cdots, a_1)$,

$$L[H_{\overline{i_1 \cdots i_{a_1}} i_j} | H_{\overline{i_1 \cdots i_{a_1}}}] \geqq \lambda_{a_1+1} \qquad (j = a_1+1, \cdots, a_1+a_2)$$

....

$$L[H_{\overline{i_1 \cdots i_{a_1}} \cdots \overline{i_{a_1+\cdots+a_{r-1}} \cdots i_{a_1+\cdots+a_r}} i_j} | H_{\overline{i_1 \cdots i_{a_1}} \cdots \overline{i_{a_1+\cdots+a_{r-1}} \cdots i_{a_1+\cdots+a_r}}}] \geqq \lambda_{a_1+\cdots+a_r+1}$$
$$(j = a_1 + \cdots + a_r + 1, \cdots, k) \tag{3-7}$$
and
$$L[H_{\overline{i_1 \cdots i_{a_1}} \cdots \overline{i_{a_1+\cdots+a_{r-1}} \cdots i_{a_1+\cdots+a_r}} 0} | H_{\overline{i_1 \cdots i_{a_1}} \cdots \overline{i_{a_1+\cdots+a_{r-1}} \cdots i_{a_1+\cdots+a_r}}}] < \lambda_{k+1} \tag{3-8}$$
hold, then we select $D_{\overline{i_1 \cdots i_{a_1}} \cdots \overline{i_{a_1+\cdots+a_{r-1}} \cdots i_{a_1+\cdots+a_r}} 0}$

(2.) if (3-7) and

$$L[H_{\overline{i_1\cdots i_{a_1}}\cdots\overline{i_{a_1+\cdots+a_{r-1}}\cdots i_{a_1+\cdots+a_r}}0} \mid H_{\overline{i_1\cdots i_{a_1}}\cdots\overline{i_{a_1+\cdots+a_{r-1}}\cdots i_{a_1+\cdots+a_r}}0}] \geqq \lambda \tag{3-9}$$

hold, then we select $D_{\overline{i_1\cdots i_{a_1}}\cdots\overline{i_{a_1+\cdots+a_{r-1}}\cdots i_{a_1+\cdots+a_r} i_{a_1+\cdots+a_{r+1}}\cdots i_k}$.

An example.

Let x_i $(i=1,2,3)$ be normally independently distributed with mean θ_i $(i=1,2,3)$ and common variance 1. For example, the rejection of a hypothesis $H_1: \infty > \theta_1 \geqq \theta_2, \theta_3 > -\infty$ when $L[H_1|G]$ is small, is equivalent to the rejection when $-2\ln L[H_1|G]$ is large.

$$-2\ln L[H_1|G] = \begin{cases} 0 & X_1 > X_2, X_3 \\ (X_2 - X_1)^2/4\,, & X_2 > X_1 > X_3 \\ (X_3 - X_1)^2/4\,, & X_3 > X_1 > X_2 \\ \sum_{r=1}^{3}(X_r - \bar{X})^2\,, & X_2, X_3 > X_1 \end{cases}$$

where $\bar{X} = \sum_{r=1}^{3} X_r/3$.

If $-2\ln L[H_1|G] \geqq \lambda_1$, H_1 is rejected, where λ_1 satisfies

$$Pr\{-2\ln L[H_1|G] \geqq \lambda_1 | H_1\} < \alpha_1\,.$$

We wish to find λ_1 under the condition where α_1 is a preassigned value in (0, 1). Putting

$$T_{21} = X_2 - X_1\,, \qquad T_{31} = X_3 - X_1\,,$$
$$\delta_{12} = \theta_1 - \theta_2 \ (>0)\,, \qquad \delta_{13} = \theta_1 - \theta_3 \ (>0)\,,$$

the joint probability density function is given by

$$f(T_{21}, T_{31}) = \frac{1}{\sqrt{3}\pi} \exp\left[-\frac{2}{3}\{(T_{21}+\delta_{12})^2 - (T_{21}+\delta_{12})(T_{31}+\delta_{13}) + (T_{31}+\delta_{13})^2\}\right].$$

When H_1 is true, the probability of rejecting H_1 is given by

$$Pr\{-2\ln L[H_1|G] \geqq \lambda_1 | H_1\} = P_1 + P_2 + P_3\,,$$

where

$$\begin{aligned} P_1 &= Pr\{(X_2 > X_1 > X_3) \cap (X_3 - X_1 \geqq c_1) | H_1\} \\ &= Pr\{(0 > T_{31}) \cap (T_{21} \geqq c_1) | H_1\}\,, \\ P_2 &= Pr\{(0 > T_{21}) \cap (T_{31} \geqq c_1) | H_1\}\,, \\ P_3 &= Pr\left\{(X_2, X_3 > X_1) \cap \left(\sum_{r=1}^{3}(X_r - \bar{X})^2 \geqq c_1^2/4\right) \middle| H_1\right\} \\ &= Pr\left\{(T_{21}, T_{31} > 0) \cap \left(T_{21}^2 + T_{21}T_{31} + T_{31}^2 \geqq \frac{3}{8}c_1^2\right) \middle| H_1\right\}, \end{aligned}$$

and

$$c_1 = 2\sqrt{\lambda_1} \quad (>0)\,.$$

We can easily show that

$$P_1 = \int_{c_1+\delta_{12}}^{\infty} \phi(t)\Phi((2\delta_{13} - t)/\sqrt{3})dt\,,$$

$$P_2=\int_{c_1+\delta_{13}}^{\infty}\phi(t)\Phi((2\delta_{12}-t)/\sqrt{3})dt,$$

and

$$P_3=\int_{\sqrt{\frac{3}{8}}c_1+\delta_{13}}^{\infty}\phi(t)[1-\Phi((2\delta_{12}-t)/\sqrt{3}+c_1/\sqrt{2})]dt$$

$$+\int_{\delta_{12}}^{\sqrt{\frac{3}{8}}c_1+\delta_{12}}\phi(t)[1-\Phi(\sqrt{c_1^2/2-(\delta_{12}-t)^2}+(2\delta_{13}-\delta_{12})/\sqrt{3})]dt$$

$$+\int_{\delta_{13}}^{\sqrt{\frac{3}{8}}c_1+\delta_{13}}\phi(t)[1-\Phi(\sqrt{c_1^2/2-(\delta_{13}-t)^2}+(2\delta_{12}-\delta_{13})/\sqrt{3})]dt,$$

where

$$\phi(t)=\frac{1}{\sqrt{2\pi}}e^{-t^2/2},\qquad \Phi(t)=\int_{-\infty}^{t}\phi(x)dx.$$

Thus the successive multiple decision procedure in our present example is summarized in the following: Let $x_i>x_j>x_k$.

$(1_\circ)$ if $x_i-x_j\geqq c_1$ and

$(1_\circ\text{-}1)$ $x_j-x_k<c_2$, select $D_{\bar{i}0}$,

$(1_\circ\text{-}2)$ $x_j-x_k\geqq c_2$, select $D_{\bar{i}\bar{j}0}$,

$(2_\circ)$ if $x_i-x_j<c_1$ and

$(2_\circ\text{-}1)$ $\sum_{r=1}^{3}(x_r-\bar{x})^2\geqq\lambda_1$, select $D_{\overline{ij}0}$,

$(2_\circ\text{-}2)$ $\sum_{r=1}^{3}(x_r-\bar{x})^2<\lambda_1$, select D_0

where c_1, c_2 and λ_1 are determined by

$$2\int_{c_1}^{\infty}\phi(t)\Phi(-t/\sqrt{3})dt+\int_{\sqrt{\frac{3}{8}}c_1}^{\infty}\phi(t)[1-\Phi(-t/\sqrt{3}+c_1/\sqrt{2})]dt$$

$$+2\int_{0}^{\sqrt{\frac{3}{8}}c_1}\phi(t)[1-\Phi(\sqrt{c_1^2/2-t^2})]dt=\alpha_1,$$

$$1-\Phi(c_2/\sqrt{2})=\alpha_2,$$

and $$\lambda_1=c_1^2/4.$$

4. Slippage problem in linear hypotheses model.

We assumed in Section 2 and 3 that the likelihood function is defined by (2-1). Under this assumption, our successive multiple decision procedure would be useful for selection and ranking problem. However there are many cases which do not satisfy (2-1) as in linear hypotheses model. In such a situation it is clear that the condition (C1-1) is not satisfied.

In this section we consider a slippage problem in a linear model, and in the next section, using the results derived in this section, we propose two successive multiple decision procedures which satisfy the condition (C1-3) under the hypotheses (1-10).

Let $Y=(Y_1, Y_2, \cdots, Y_n)'$ be a random vector distributed according to the multivariate normal distribution with mean vector ξ and covariance matrix $\sigma^2 I$. About ξ_i $(i=1, 2, \cdots, n)$ we suppose they are divided into $k+1$ categories consisting of m except for the last category consisting $n-km$ for which assume $\xi_i=0$, $i=km+1, \cdots, n$.

We consider the problem of jointly testing the k hypotheses

$$H_0^{(i)}: \xi_{(i-1)m+1}=\cdots=\xi_{im}=0 \qquad (i=1, 2, \cdots, k).$$

Now we impose the requirement that

(C4-1) the test procedure should be invariant under the group $G_{11}, \cdots, G_{1k}$ and G_2 of transformations:

G_{1j} $(j=1, 2, \cdots, k)$: all orthognal transformations of $Y_{(j-1)m+1}, \cdots, Y_{jm}$.

G_2: $Y_i'=cY_i$, $i=1, 2, \cdots, n$; $c\neq 0$.

It follows that under the condition (C4-1), the problem of jointly testing k hypotheses $H_1^{(0)}, \cdots, H_0^{(k)}$ reduces to that of jointly testing k hypotheses $H_0^{(i)}$: $\lambda_i=0$ $(i=1, 2, \cdots, k)$ subject to the k statistics $U_1/U, \cdots, U_k/U$. Here $\lambda_i=\xi^{(i)}\xi^{(i)}/2\sigma^2$ $(i=1, 2, \cdots, k;\ km=r)$. It is well known that U_i/σ^2 $(i=1, 2, \cdots, k)$ be distributed according to the noncentral χ^2 distributions with the common m degrees of freedom and noncentrality parameters λ_i, and U/σ^2 has the central χ^2-distribution with $\nu=n-r$ degrees of freedom, and $U_1, \cdots, U_k$ and U are mutually independent. We consider a multiple decision problem with respect to testing hypotheses for k unknown parameters $\lambda_1, \cdots, \lambda_k$.

There is the k sample slippage problem as an important class of multiple decision problems. The classical slippage problems are to find a satistical procedure which will, on the basis of the observation, decide if all the k parameters are eqnal to each other, and if not, which one out of k parameters has slipped to the right. Hence $k+1$ decision corresponding to $k+1$ all possible hypotheses are introduced. We now return to the linear model given in the beginning of this section, and introduce a slippage type hypotheses as follows.

$$H_0: \lambda_1=\cdots=\lambda_k=0$$

$$H_i: \lambda_i=\Delta, \qquad \lambda_1=\cdots=\lambda_{i-1}=\lambda_{i+1}=\cdots=\lambda_k=0 \qquad (i=1, 2, \cdots, k)$$

where $\Delta>0$.

The slippage problem is to find an optimum decision procedure in the sense that maximizes the probability of making the correct decision subject to the following conditions

(C4-2) $$Pr\{D_0|H_0\}\geqq 1-\alpha \quad \text{where} \quad \alpha\in(0, 1),$$

(C4-3) the decision precedure should be symmetric, that is,

$$Pr\{D_1|H_1\}=\cdots=Pr\{D_k|H_k\}.$$

This type of problem has been first investigated by Paulson [9], Truax [12] for the mean and the variance of k normal populations and in more general formulation by Hall, Kudo and Yeh [4], [5], [14] among other authors. According their procedure, we have the following theorem.

Put $U_{i_1}=\mathrm{Max}(U_1, \cdots, U_k)$, and set out the quantity $S_{(k)}=U+\sum_{j=1}^{k} U_j$. Let D_{i_1} be the decision corresponding to distribution yielding U_{i_1}.

THEOREM 4-1. *The following decision procedure for the splippage problem is optimum under the conditions* (C4-1), (C4-2), *and* (C4-3).

$$\begin{aligned} &\text{If} \quad U_{i_1}/S_{(k)} > L(\alpha\,;\,\nu, k, m)\,, \qquad \text{select} \quad D_{i_1}\,,\\ &\text{If} \quad U_{i_1}/S_{(k)} \leqq L(\alpha\,;\,\nu, k, m)\,, \qquad \text{select} \quad D_0\,, \end{aligned} \tag{4-1}$$

where $L(\alpha\,;\,\nu, k, m)$ *is a constant whose precise value is determined by condition* (C4-2) *and does not depend on* Δ.

PROOF. Because of the condition (C4-1), it is clear that $U, U_1, \cdots, U_k$ form the maximal invariant statistics, and by applying Theorem 1 and 2 of Hall and Kudo [4] it is straightforward to prove the theorem.

Next we derive an approximative formula for $L(\alpha\,;\,\nu, k, m)$, and this is obtained in a manner similar to the calculation given by Cochran [2]. Put

$$W_j=U_j/\Big(U+\sum_{i=1}^{k} U_i\Big)\,, \qquad j=1, 2, \cdots, k\,.$$

The joint probability density function of $W_1, \cdots, W_k$ is given by

$$\frac{e^{-\sum_{i=1}^{k}\lambda_i}}{\Gamma(\nu/2)}\Big(1-\sum_{i=1}^{k} w_i\Big)^{\frac{\nu}{2}-1}\cdot\prod_{i=1}^{k}\sum_{m_i=0}^{\infty}\frac{\lambda_i^{m_i}}{m_i\,!}\,\frac{\Gamma\Big[\frac{1}{2}(\nu+km)+\sum_{i=1}^{k} m_i\Big]}{\Gamma\Big(\frac{m}{2}+m_i\Big)}\,w_i^{\frac{m}{2}+m_i-1}\,. \tag{4-2}$$

The probability that the largest of the ratios $W_1, \cdots, W_k$ exeeds L is given by

$$Pr\{\mathrm{Max}\, W_j > L\,|\,H_0\} = \sum_{i=1}^{k}(-1)^{i-1}\binom{k}{i}P_i(L\,|\,H_0) \tag{4-3}$$

where $P_1(L\,|\,H_0)$ is the probability that a specified one of the ratios exeeds L when H_0 is correct, $P_2(L\,|\,H_0)$ the probability that a specified pair of the ratios both exeed L, and so on. Clearly

$$P_k(L\,|\,H_0)=0 \qquad \text{when} \quad L > 1/k\,, \tag{4-4}$$

so that the number of non-zero terms in (4-3) is the greatest integer less than $1/L$.

When D_0 is correct, we put $U_j/\sigma^2=\chi_j^2$ and $U+\sum_{\substack{i=1\\ i\neq j}}^{k} U_i/\sigma^2=\chi_{(j)}^2$, and hence these statistics are independently distributed according to the χ^2-distributions with m and $\nu+(k-1)m$ degrees of freedom, respectively. Furthermore making the change of variables $\chi_j^2=2W_jX$ and $\chi_j^2+\chi_{(j)}^2=2X$, the joint probability density function of W_j, and X is given by

$$\Big[\Gamma\Big(\frac{m}{2}\Big)\Gamma\Big(\frac{\nu+(k-1)m}{2}\Big)\Big]^{-1} w_j^{\frac{m}{2}-1}(1-w_j)^{\frac{\nu+(k-1)m}{2}-1}x^{\frac{\nu+km}{2}-1}e^{-\frac{x}{2}}\,.$$

Integrating from 0 to ∞ with respect to x, we obtain the probability density function of W_j as follows.

$$w_j^{\frac{m}{2}-1}(1-w_j)^{\frac{\nu+(k-1)m}{2}-1}\Big/B\Big(\frac{m}{2},\ \frac{\nu+(k-1)m}{2}\Big). \tag{4-5}$$

Thus we have

$$P_1(L\,|\,H_0)=I_{1-L}\Big(\frac{\nu+(k-1)m}{2},\ \frac{m}{2}\Big),\qquad 0\leqq L\leqq 1. \tag{4-6}$$

Since $P_2(L\,|\,H_0)<[P_1(L\,|\,H_0)]^2$ for all values of ν, k and m, the upper α level of significance of L can be evaluated perhaps to a good approximation using only the first term of (4-3).

Hence, if $\tilde{L}_\alpha$ is determined from

$$P_1(\tilde{L}_\alpha\,|\,H_0)=\alpha/k\,, \tag{4-7}$$

it follows from (4-3) that

$$\alpha\Big(1-\frac{\alpha}{2}\Big)<\alpha\Big(1-\frac{k-1}{2k}\alpha\Big)<P(\tilde{L}_\alpha\,|\,H_0)\leqq\alpha \tag{4-8}$$

and this interval is sufficiently narrow when α is small, so that a good approximate value of $L(\alpha\,;\,\nu,\,km)$ is given by $\tilde{L}_\alpha$.

From (4-6) it follows that

$$\tilde{L}_\alpha=L(\alpha\,;\,\nu,\,k,\,m)=1-x_{\alpha/k}\,,$$

where

$$x_{\alpha/k}=x_{\alpha/k}\Big[\frac{m}{2},\ \frac{\nu+(k-1)m}{2}\Big]$$

is the lower α/k level of significance of the beta distribution. In summary we have that $L(\alpha,\,\nu,\,k,\,m)$ is approximately given by

$$L(\alpha\,;\,\nu,\,k,\,m)\doteqdot\tilde{L}_\alpha=F_{\alpha/k}\Big/\Big(F_{\alpha/k}+\frac{\nu}{m}+k-1\Big)\,, \tag{4-9}$$

where

$$F_{\alpha/k}=F^{m}_{\nu+(k-1)m}(\alpha/k)$$

is the upper α/k level of significance of the F-distribution with m and $\nu+(k-1)m$ degrees of freedom. And the degree of approximation is eveluated by (4-8).

In the next section, we are required that the condition $L(\alpha\,;\,\nu,\,k,\,m)>1/2$, so from (4-6) and (4-7) we have

$$0<\alpha<kP_1(1/2\,|\,H_0)=kI_{1/2}\Big(\frac{\nu+(k-1)m}{2},\ \frac{m}{2}\Big)$$

when $1/2<\tilde{L}_\alpha<1$. Therefore the values of $kP_1(1/2\,|\,H_0)$ for some values of ν, k and m are tabulated in the Table. When α is a very small positive number on (0, 1), it is seen from the Table that the condition $L(\alpha\,;\,\nu,\,k,\,m)>1/2$ is satisfied, if ν is sufficiently large.

Table. The value of $kP_1(1/2|H_0)$

1. $k=2$

ν \ m	1	2	3	5	10	20	30
1	.58579	.70711	.76256	.81781	.87238	.90826	.92777
2	.36338	.50000	.57559	.66047	.75391	.82380	.85554
5	.09965	.17678	.23776	.32839	.46289	.59355	.65836
10	.01374	.03125	.05049	.09006	.17957	.30746	.39153
20	.00032	.00098	.00211	.00544	.01921	.06143	.10813
30	.00001	.00003	.00008	.00026	.00154	.00904	.02263
50					.00001	.00011	.00054

2. $k=3$

ν \ m	1	2	3	5	10	20	30
1	.54507	.53033	.48141	.38440	.21929	.07815	.02904
2	.34835	.37500	.35664	.29818	.17770	.06416	.02414
5	.09944	.13258	.14188	.13509	.09439	.03713	.01447
10	.01405	.02344	.02913	.13551	.02882	.01356	.00570
20	.00034	.00073	.00112	.00175	.00232	.00160	.00080
30	.00001	.00002	.00004	.00008	.00016	.00016	.00010

3. $k=5$

ν \ m	1	2	3	5	10	20
1	.37793	.22097	.12623	.04197	.00307	.00002
2	.24913	.15625	.09195	.03143	.00228	.00001
5	.07478	.05524	.03521	.01371	.00106	.00001
10	.01095	.00977	.00694	.00291	.00026	
20	.00027	.00031	.00026	.00013	.00002	
30	.00001	.00001	.00001	.00001		

4. $k=10$

ν \ m	1	2	3	5
1	.10120	.01381	.00193	.00004
2	.06872	.00977	.00046	.00003
5	.02189	.00366	.00051	.00001
10	.00338	.00061	.00010	

5. $k=20$

ν \ m	1	2
1	.00676	.00003
2	.00323	.00001
5	.00019	
10	.00001	

5. Successive multiple decision procedures.

Using the results in Section 4, we propose and consider an algorithm related with successive multiple decision procedure for testing hypotheses (1-10) stated in Section 1. The parameters θ_i $(i=1, \cdots, k)$ and θ_0 in Section 1 correspond to λ_i $(i=1, \cdots, k)$ and 0 in Section 4, respectively. We arrange the observations $u_1, \cdots, u_k$ in decending order of magnitude as $u_{i_1} \geqq u_{i_2} \geqq \cdots \geqq u_{i_k}$ and set out the quantity $s_{(k-a+1)} = u + \sum_{j=1}^{k} u_{i_j}$. Our algorithm is as follows.

Procedure I.

(1_0) if $u_{i_1}/s_{(k)} \leqq L(\alpha_1;\ \nu, k, m)$, select D_0

(2_0) if $u_{i_1}/s_{(k)} > L(\alpha_1;\ \nu, k, m)$

and $u_{i_2}/s_{(k-1)} \leqq L(\alpha_2;\ \nu, k-1, m)$, select $D_{i_1 0}$

....

(r_0) if $u_{i_1}/s_{(k)} > L(\alpha_1;\ \nu, k-1, m)$

$u_{i_2}/s_{(k-1)} > L(\alpha_2;\ \nu, k-1, m)$

....

$u_{i_{r-1}}/s_{(k-r+2)} > L(\alpha_{r-1};\ \nu, k-r+2, m)$

and $u_{i_r}/s_{(k-r+1)} \leqq L(\alpha_r;\ \nu, k-r+1, m)$, select $D_{i_1 i_2 \cdots i_{r-1} 0}$

....

$(m+1_0)$ if $u_{i_1}/s_{(k)} > L(\alpha_1;\ \nu, k, m)$

$u_{i_2}/s_{(k-1)} > L(\alpha_2;\ \nu, k-1, m)$

....

and $u_{i_k}/s_{(1)} > L(\alpha_k;\ \nu, k-1, m)$, select $D_{i_1 i_2 \cdots i_k}$.

We will show that Procedure I has some desirable properties. We at first prepare with some Lemmas for this purpose.

LEMMA 5-1. *Let A_{NM}, A_{MN}, T_1, T_2 be some positive constants. If $A_{NM} > A_{MN}$ for $M > N$ and $T_2 > T_1$, then*

$$\sum_{N=0}^{\infty} T_1^N \sum_{M=0}^{\infty} T_2^M A_{NM} > \sum_{N=0}^{\infty} T_2^N \sum_{M=0}^{\infty} T_1^M A_{MN}$$

holds.

The proof of this Lemma is simple and it omited. Consider only a hypothesis $H_{12\cdots r}$ for the sake of simplicity.

LEMMA 5-2. *When $\Delta_j \geqq \Delta_i \geqq 0$ and $L(\alpha;\ \nu, k, m) > 1/2$, the slippage test* (4-1) *satisfies*

$$Pr\{D_j | H_{12\cdots r0}\} \geqq Pr\{D_i | H_{12\cdots r0}\} \tag{5-1}$$

for $i, j = 1, 2, \cdots, k$ and the equality holds when $\Delta_j = \Delta_i$.

PROOF. Putting $L = L(\alpha;\ \nu, k, m)$, we have

$$Pr\{D_j | H_{12\cdots k}\} = Pr\Big\{U_j > U_1, \cdots, U_{j-1}, U_{j+1}, \cdots, U_k \text{ and } U_j/\Big(U + \sum_{i=1}^{k} U_i\Big) > L | H_{12\cdots k}\Big\}.$$

It follows from $L > 1/2$ that

$$Pr\{D_j | H_{12\cdots k}\}$$

$$= Pr\Big\{U_j > \frac{L}{1-L}\Big(U + \sum_{\substack{i=1 \\ i\neq j}}^{k}{}' U_i\Big)\Big| H_{12\cdots k}\Big\}$$

$$= \frac{e^{-\sum_{i=1}^{k}\Delta_i}}{\Gamma(\nu/2)}\int_0^\infty \cdots \int_0^\infty t^{\frac{\nu}{2}-1}e^{-t}\Big[\prod_{\substack{i=1 \\ i\neq j}}^{k}{}' \sum_{m_i=0}^{\infty}\frac{\Delta_i^{m_i}}{m_i!}\frac{1}{\Gamma\big(\frac{m}{2}+m_i\big)}t^{\frac{m}{2}+m_i-1}e^{-t_i}$$

$$\cdot\Big(\sum_{m_j=0}^{\infty}\frac{\Delta_j^{m_j}}{m_j!}\frac{1}{\Gamma\big(\frac{m}{2}+m_j\big)}\int_{\frac{L}{1-L}(t+\sum'_{i\neq j} t_i)}^{\infty} t_j^{\frac{m}{2}+m_j-1}e^{-t_j}dt_j\Big)dt_i\Big]dt$$

$$= \frac{e^{-\sum_{i=1}^{k}\Delta_i}}{\Gamma(\nu/2)}\int_0^\infty \cdots \int_0^\infty t^{\frac{\nu}{2}-1}e^{-\frac{t}{1-L}}\Big[\prod_{\substack{i=1 \\ i=j,1}}^{k}{}'' \sum_{m_i=0}^{\infty}\frac{\Delta_i^{m_i}}{m_i!}\frac{1}{\Gamma\big(\frac{m}{2}+m_i\big)}t_i^{\frac{k}{2}+m_i-1}e^{-\frac{t_i}{1-L}}$$

$$\cdot\sum_{m_l=0}^{\infty}\frac{\Delta_l^{m_l}}{m_l!}\frac{1}{\Gamma\big(\frac{m}{2}+m_l\big)}\sum_{m_j=0}^{\infty}\frac{\Delta_j^{m_j}}{m_j!}\frac{1}{\Gamma\big(\frac{m}{2}+m_j\big)}\sum_{s=0}^{\frac{m}{2}+m_j-1}\frac{\Gamma\big(\frac{m}{2}+m_j\big)}{\Gamma(s-1)}\Big(\frac{L}{1-L}\Big)^s$$

$$\cdot\sum_{i=1}^{s}\binom{s}{i}\Big(t+\sum_{\substack{i=1 \\ i\neq j,l}}^{k}{}'' t_i\Big)^{s-1}(1-L)^{\frac{m}{2}+m_l+i}\Gamma\Big(\frac{m}{2}+m_l+i\Big)dt_i\Big]dt.$$

We here put

$$B_{MN} = \Gamma\Big(\frac{m}{2}+N\Big)\sum_{s=0}^{\frac{m}{2}+N-1}\sum_{i=0}^{s}(1-L)^{\frac{m}{2}+M+i}\Gamma\Big(\frac{m}{2}+M+i\Big).$$

When $M > N$, we have

$$B_{NM} - B_{MN} = \sum_{s=m/2+N}^{m/2+M-1}\Gamma\Big(\frac{m}{2}+M\Big)\sum_{i=0}^{s}(1-L)^{\frac{m}{2}+N+i}\Gamma\Big(\frac{m}{2}+N+i\Big)$$

$$+\sum_{s=0}^{\frac{m}{2}+N-1}\Gamma\Big(\frac{m}{2}+N\Big)\sum_{i=0}^{s}(1-L)^{\frac{m}{2}+N+i}\Gamma\Big(\frac{m}{2}+N+i\Big)$$

$$\cdot\Big[\Big(\frac{m}{2}+N\Big)\cdots\Big(\frac{m}{2}+M-1\Big)-(1-L)^{M-N}\Big(\frac{m}{2}+N+i\Big)\cdots\Big(\frac{m}{2}+N+i\Big)\Big].$$

Using max $i = \frac{m}{2}+N-1$ in the last term of the right hand side, the inside of the brackets becomes

$$\Big[\Big(\frac{m}{2}+N\Big)\cdots\Big(\frac{m}{2}+M-1\Big)-(1-L)^{M-N}(m+2N-1)\cdots(m+M+N-2)\Big]$$

$$= \Big(\frac{m}{2}+N\Big)\cdots\Big(\frac{m}{2}+M-1\Big)\{2(1-L)\}^{M-N}\Big(\frac{m}{2}+N-1\Big)\cdots\Big(\frac{m}{2}+\frac{M+N}{2}-1\Big).$$

Since $L > 1/2$ and thus $2(1-L) < 1$, the right hand side clearly takes the positive value. Therefore $B_{NM} > B_{MN}$ for $M > N$. From Lemma 5-1, we know that

$$Pr\{D_j|H_{12\cdots k}\} > Pr\{D_i|H_{12\cdots k}\} \qquad \text{for} \quad \Delta_j > \Delta_i \tag{5-2}$$

and

$$Pr\{D_j|H_{12\cdots k}\} = Pr\{D_i|H_{12\cdots k}\} \qquad \text{for} \quad \Delta_j = \Delta_i\,. \tag{5-3}$$

Putting $\Delta_{s+1} = \cdots = \Delta_k = 0$ and using (5-2) and (5-3), the proof is complete.

From Theorem 4-1 and Lemma 5-2, we have the following.

THEOREM 5-1. *In each step of Procedure I, the condition* (C1-3) *is satisfied.*

It is easily seen from Theorem 4-1 that if H_{i_10} $(i=1, \cdots, k)$ is true, then Procedure I is optimum in the sense that it maximizes the probability of correct decision. However, it is shown in the following section that if $H_{i_1\cdots i_r0}$ $(r>1)$ is true, then Procedure I attaines to the maximum: $\underset{\Phi_{i_r0(i_r\neq i_1,\cdots,i_{r-1})}}{\text{Max}} Pr\{D_{(i_1\cdots i_{r-1})i_r0}|H_{i_1\cdots i_r0}\}$, but in general does not attain to the maximum of $Pr\{D_{i_1\cdots i_{r-1}}|H_{i_1\cdots i_r0}\}$.

We next propose and consider another algorithm related with successive multiple decision procedure which is perhaps more robust against the erroneous decisions previously made in each steps. It is as follows:

Procedure II

(1_0) if $u_{i_1}/u \leqq S(\alpha_1\,;\, \nu, k, m)$, select D_0,

(2_0) if $u_{i_1}/u > S(\alpha_1\,;\, \nu, k, m)$,

and $u_{i_2}/u \leqq S(\alpha_2\,;\, \nu, k-1, m)$, select D_{i_10},

....

(r_0) if $u_{i_1}/u > S(\alpha_1\,;\, \nu, k, m)$

$u_{i_2}/u > S(\alpha_2\,;\, \nu, k-1, m)$

....

$u_{i_{r-1}}/u > S(\alpha_{r-1}\,;\, \nu, k-r+2, m)$

and $u_{i_r}/u \leqq S(\alpha_r\,;\, \nu, k-r+1, m)$, select $D_{i_1\cdots i_{r-1}0}$,

....

$(k+1_0)$ if $u_{i_1}/u > S(\alpha_1\,;\, \nu, k, m)$

$u_{i_2}/u \leqq S(\alpha_2\,;\, \nu, k-1, m)$

....

and $u_{i_k}/u > S(\alpha_k\,;\, \nu, k, m)$, select $D_{i_1\cdots i_k}$,

where $S(\alpha_i\,;\, \nu, k, m)$ is a constant whose precise value is determined by condition (C4-2).

According to similar method to that of proof in Theorem 5-1, we have the following.

THEOREM 5-2. *In each step of Procedure II, the condition* (C1-3) *is satisfied.*

REMARK. The general aspect of Procedure I and II in a linear hypotheses model can be observed in terms of the general formulation and discussion of the following successive multiple decision procedure. First we constract a sequence of decision sets as follows:

$$\left.\begin{aligned}
\mathfrak{D}_0 &= (D_0, D_{10}, \cdots, D_{k0}), \\
\mathfrak{D}_{(i_1)} &= (D_{(i_1)0}, D_{(i_1)i_20}, \cdots, D_{(i_1)i_1-10} D_{(i_1)i_1+10}, \cdots, D_{(i_1)i_k0}) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\mathfrak{D}_{(i_1\cdots i_r)} &= (D_{(i_1\cdots i_r)0}, D_{(i_1\cdots i_r)\zeta 0} : \zeta = i_{r+1}, i_{r+2}, \cdots, i_k) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\mathfrak{D}_{(i_1\cdots i_{k-1})} &= (D_{(i_1\cdots i_{k-1})0}, D_{(i_1\cdots i_{k-1})i_k}),
\end{aligned}\right\} \tag{5-4}$$

where $D.$ denotes the decision to accept the corresponding hypothesis $H.$ in the following

$$\begin{aligned}
H_{(i_1\cdots i_r)0} &: \theta_{i_{r+1}} = \cdots = \theta_{i_k} = \theta_0, \\
H_{(i_1\cdots i_r)i_{r+1}0} &: \theta_{i_{r+1}} > \theta_{i_{r+2}} = \cdots = \theta_{i_k} = \theta_0 \qquad (r = 1, 2, \cdots, k-2) \\
H_{(i_1\cdots i_{k-1})i} &: \theta_{i_k} > \theta_0.
\end{aligned} \tag{5-5}$$

We now impose the requirements for selecting an element of $\mathfrak{D}_{(i_1\cdots i_r)}$ that

(C5-1) $Pr\{D_{(i_1\cdots i_r)0} | H_{(i_1\cdots i_r)0}\} \geqq 1 - \alpha_{r+1}, \ \alpha_{r+1} \in (0, 1) \qquad (r = 0, 1, \cdots, k-1)$

(C5-2) $Pr\{D_{(i_1\cdots i_r)i_{r+1}0} | H_{(i_1\cdots i_r)i_{r+1}0}\} = \cdots Pr\{D_{(i_1\cdots i_r)i_k0} | H_{(i_1\cdots i_r)i_k0}\} \qquad (r = 0, 1, \cdots, k-1)$

where

$$D_{(i_1\cdots i_r)0} = D_0$$

and

$$D_{(i_1\cdots i_r)i_{r+1}0} = \begin{cases} D_{i_10} & \text{when } r = 0, \\ D_{(i_1\cdots i_{k-1})i_k} & \text{when } r = k-1. \end{cases}$$

Let us consider the following rule of statistical procedures for selecting a decision to accept one of the hypotheses (1-10):

(1) First an element of $\mathfrak{D}_0$ is selected with the help of a procedure which satisfies the condition (C5-1) and (C5-2) with $r=0$.
  (1_1) If D_0 is selected in step (1), then we stop and hence decide D_0.
  (1_2) If D_{i_10} is selected in step (1), then further examine $\mathfrak{D}_{(i_1)}$.
(2) In the case (1_2) an element of $\mathfrak{D}_{(i_1)}$ is selected with the help of a procedure which satisfies the condition (C5-1) and (C5-2) with $r=1$.
  (2_1) If $D_{(i_1)0}$ is selected in step (2), then we stop and decide D_{i_10}.
  (2_2) If $D_{(i_1)i_20}$ is selected in step (2), then further examine $\mathfrak{D}_{(i_1i_2)}$.

We continue until an element $D_{(i_1\cdots i_r)0}$ of $\mathfrak{D}_{(i_1\cdots i_r)}$ is selected with the help of a procedure which satisfies the condition (C5-1) and (C5-2), and hence we stop and decide $D_{i_1\cdots i_r0}$.

And we shall call the procedure proposed above a local optimal procedure if the condition (C1-3) is satisfied in each of all steps. Under this definition, Theorem 5-1 and 5-2 can be rewrited as "Procedure I and II are local optimal".

It is easily seen from Theorem 4-1 that if H_{i_10} ($i_1 = 1, \cdots, k$) is true, then Procedure I is optimum in the sense that is maximized the probability of correct decision. However note that when $H_{i_1\cdots i_r0}$ ($r > 0$) is true, this local optimal procedure (and hence Procedure I too) enjoyes a kind of local optimum property that is, if the outcome of the previuos step of the test happens to be correct the present test has an optimum property.

It should be noted that in case of $m=1$, our present Procedure I and II is applicable to find a successive procedure of fitting a sequence of orthogonal functions that was suggested by T. Kitagawa [7], [8].

6. Comparison of Procedure I and II.

In this section, we compare the powers of Procedure I and II. For the sake of simplicity, we consider the case of $m=2$ and a random model instead of parametric model in the preceding sections.

We assume the mean squares U_i $(i=0, 1, 2)$ are independently distributed as $\chi_i^2\sigma_i^2/n_i$, where χ_i^2 is the central χ^2 statistics based on n_i degrees of freedom. Suppose that we are interested in testing the following multiple hypotheses

$$\begin{array}{ll} H_0 : \theta_2=\theta_1=1 & \\ H_{20}: \theta_2>\theta_1=1 & H_{10}: \theta_1>\theta_2=1 \\ H_{21}: \theta_2>\theta_1>1 & H_{12}: \theta_1>\theta_2>1 \end{array} \tag{4-1}$$

where $\theta_i=\sigma_i^2/\sigma_0^2$ $(i=1, 2)$. When the observations u_1, u_2 are arranged in decending order of magnitude as $u_{(2)}>u_{(1)}$, Procedure I and II are given by

Procedure I

$$\begin{array}{llll} 1_0) & \text{if} & u_{(2)}/(u_0+u_{(1)})\leqq L_2\,, & \text{select } D_0\,, \\ 2\text{-}1_0) & \text{if} & u_{(2)}/(u_0+u_{(1)})>L_2 & \\ & \text{and} & u_{(1)}/u_0\leqq L_1\,, & \text{select } D_{(2)0}\,, \\ 2\text{-}2_0) & \text{if} & u_{(2)}/(u_0+u_{(1)})>L_2 & \\ & \text{and} & u_{(1)}/u_0>L_1\,, & \text{select } D_{(2)(1)}\,. \end{array}$$

Procedure II

$$\begin{array}{llll} 1_0) & \text{if} & u_{(2)}/u_0\leqq S_2\,, & \text{select } D_0\,, \\ 2\text{-}1_0) & \text{if} & u_{(2)}/u_0>S_2 \text{ and } u_{(1)}/u_0\leqq S_1\,, & \text{select } D_{(2)0}\,, \\ 2\text{-}2_0) & \text{if} & u_{(2)}/u_0>S_2 \text{ and } u_{(1)}/u_0>S_1\,, & \text{select } D_{(2)(1)}\,. \end{array}$$

We now consider the simplest case as $n_1=n_2=2$. The joint probability density function of $W=U_2/U_0$ and $V=U_1/U_0$ is given by

$$f(w, v\,;\ \theta_2, \theta_1)=\frac{1}{\theta_1\theta_2}\frac{n_0}{2}\left(\frac{n_0}{2}+1\right)\left(1+\frac{v}{\theta_1}+\frac{w}{\theta_2}\right)^{-\frac{n_0}{2}-2}.$$

In the 1_0 step, let $\alpha_{\rm I}$ and $\alpha_{\rm II}$ be the sizes of Procedure I and II, $\beta_{\rm I}(\theta_2, \theta_1)$ and $\beta_{\rm II}(\theta_2, \theta_1)$ denote the powers of Procedure I and II assumed $\theta_2>\theta_1$, respectively. Hence we have

$$1-\alpha_{\mathrm{I}}=Pr\{V\leqq W\leqq L_2(1+V)\mid H_0\}+Pr\{W\leqq V\leqq L_2(1+W)\mid H_0\}$$

$$=\begin{cases} 2\int_0^\infty\int_w^{L_2(1+w)}f(w,v;\,1,1)dv\,dw & \text{for } L_2>1,\\ 2\left(\int_0^{\frac{L_2}{1-L_2}}\int_{L_2(1+v)}^\infty+\int_{\frac{L_2}{1-L_2}}^\infty\int_v^\infty\right)f(w,v;\,1,1)dv\,dw & \text{for } L_2<1,\end{cases}$$

and then

$$\alpha_{\mathrm{I}}=\begin{cases} 2/(1+L_2)^{\frac{n_0}{2}+1} & \text{for } L_2>1\\ 2[1+L_2(1-L_2)^{\frac{n_0}{2}}]/(1+L_2)^{\frac{n_0}{2}+1} & \text{for } L_2<1.\end{cases}$$

Similarly, we have

$$\alpha_{\mathrm{II}}=2(1+S_2)^{-n_0/2}-(1+2S_2)^{-n_0/2},$$

$$\beta_{\mathrm{I}}(\theta_2,\theta_1)=\begin{cases}\left(1+\frac{\theta_1}{\theta_2}L_2\right)^{-1}\left(1+\frac{1}{\theta_2}L_2\right)^{-n_0/2} & \text{for } L_2>1,\\ 1-\left(1+\frac{\theta_1}{\theta_2}\right)^{-1}\Big[\Big\{1+\frac{1}{\theta_2}+\frac{1}{\theta_1}\left(1+\frac{\theta_1}{\theta_2}\right)\left(\frac{L_2}{1-L_2}\right)\Big\}^{-n_0/2} & \\ \qquad -1\Big\{1+\frac{1}{\theta_1}\left(1+\frac{\theta_1}{\theta_2}\right)\left(\frac{L_2}{1-L_2}\right)\Big\}^{-n_0/2}\Big] & \text{for } L_2<1\end{cases}$$

and

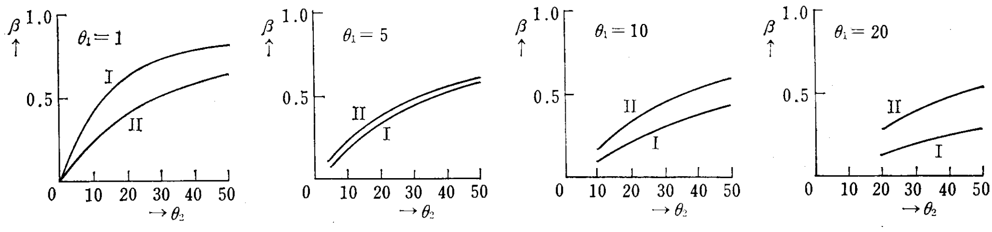
$$\beta_{\mathrm{II}}(\theta_2,\theta_1)=\left(1+\frac{\theta_1}{\theta_2}S_2\right)^{-n_0/2}-\left(1+\frac{\theta_2}{\theta_1}\right)^{-1}\left[1+\left(1+\frac{\theta_2}{\theta_1}\right)\frac{1}{\theta_2}S_2\right]^{-n_0/2}.$$

As an example, we consider the case of $n_0=2$ and $\alpha_{\mathrm{I}}=\alpha_{\mathrm{II}}=0.05$. We obtain $L_2=5.32455$ and $S_2=28.82952$. If the hypothesis H_{20} is true, then it should be satisfied according to Theorem 4-1 that

$$\beta_{\mathrm{I}}(\theta_2,1)\geqq\beta_{\mathrm{II}}(\theta_2,1).$$

It is easily shown from the above results that this statement is correct and the inequality holds strictly. However, when the hypothesis H_{21} is true, we cannot expect the relation that the power of Procedure I is better than that of Procedure II for all possible values of θ_2 and θ_1. These relations is clearly shown in Figure.

Fig. Power of Procedure I and II ($k=m=2$)



Consequently, it was observed that in our successive multiple decision problem there does not exist an overall optimum procedure in the sense that jointly maximizes the probabilities of correct decisions for all possible hypotheses under the condition (C5-1), (C5-2) and (C1-3).

By considering of the present example, the following procedure may be suggested in case of successive multiple decision problem in linear model. First we apply Procedure II, that is, if

$$u_{i_j}/u > S(\alpha_j\,;\ \nu,\ k-j+1,\ m) \qquad (j=1, \cdots, r-1) \tag{5-1}$$

and

$$u_{i_r}/u \leqq S(\alpha_r\,;\ \nu,\ k-r+1,\ m)$$

hold, and further we apply the slippage test, that is, if

$$u_{i_r}/s_{(k-r+1)} \leqq L(\alpha_r\,;\ \nu,\ k-r+1,\ m) \tag{5-2}$$

holds, then we select $D_{i_1\cdots i_{r-1}0}$. And if (5-2) does not hold, then we return to (5-1) with $j=r+1$, and so on. Of course, such an algorithm as proposed above involves a large scale table which has to be evaluated by numerical computation.

Acknowledgements

The author is deeply indebted to Professor A. Kudô, who went through the original manuscript and improved the exposition very thoroughly. And the auther is gratful for the encouragements and criticism given by Professor T. Kitagawa and Professor A. Kudo while this paper is being worked out.

References

[1] BRUNK, H. D. (1955). *Maximum likelihood estimation of monotone parameters.* Ann. Math. Statist. **26**, 607-616.

[2] COCHRAN, W. G. (1941). *The distribution of the largest of a set of estimated variances as a fraction of their total.* Ann. Eugen., London. **11**, 47-52.

[3] DARROCH, J. N. and SIVELY, S. D. (1963). *On testing more than one hypothesis.* Ann. Math. Statist. **34**, 555-567.

[4] HALL, I. J. and KUDÔ, A. (1968). *On slippage tests—(I), A generalization of Neyman Pearson's lemma.* Ann. Math. Statist. **39**, 1693-1699.

[5] HALL, I. J. and KUDÔ, A. (1968). *On slippage tests—(II), Similar slippage tests.* Ann. Math. Statist. **39**, 2029-2037.

[6] HOGG, R. V. (1961). *On the resolution of statistical hypotheses.* J. Amer. Statist. Assoc. **56**, 978-989.

[7] KITAGAWA, T. (1956). *Some aspects of stochastically approximative analysis.* Bull. Math. Statist. **6**, 109-120.

[8] KITAGAWA, T. (1963). *Estimation after preliminary tests of significance.* Univ. California Publications in Statist. **3**, 147-186.

[9] PAULSON, E. (1952). *An optimum solution to the k-sample slippage problem for the normal distribution.* Ann. Math. Statist. **23**, 610-616.

[10] ROBERTSON, T. and WALTMAN, P. (1968). *On estimating monotone parameters.* Ann. Math. Statist. **39**, 1030-1039.

[11] SEBER, G. A. F. (1964). *Liner hypotheses and induced tests.* Biometrika. **51**, 41-47.

[12] TRUAX, D. R. (1953). *An optimum slippage test for the variance of k normal distributions.* Ann. Math. Statost. **24**, 669-674.

[13] VAN EDEN, C. (157). *Maximum likelihood estimation of partially or completely ordered parameters,* I and II. Indag. Math. **19**, 128-136, 201-211.

[14] YEH, N. C. and KUDÔ, A. (1970). *On slippage tests—(III), The least favorable distribution.* Tamkang J. Math. **1**, 57-64.