

A CLOSED SEQUENTIAL PROCEDURE SELECTING THE BEST POPULATION IN A FAMILY OF POPULATIONS WITH ONE PARAMETER EXPONENTIAL DISTRIBUTIONS

Nomachi, Yukio
Miyazaki University

<https://doi.org/10.5109/13023>

出版情報：統計数理研究. 12 (3/4), pp.21-34, 1967-03. 統計科学研究会
バージョン：
権利関係：



A CLOSED SEQUENTIAL PROCEDURE SELECTING THE BEST POPULATION IN A FAMILY OF POPULATIONS WITH ONE PARAMETER EXPONENTIAL DISTRIBUTIONS

By

Yukio NOMACHI

(Miyazaki University)

§1. Summary.

The problem to select the best population in some specified sense from several assigned populations in the light of samples drawn from them is very important in practical situations. An experimenter who is faced with this problem may select the best population according to a certain statistical procedure in view of the informations supplied by samples drawn from these populations.

In this connection several different statistical procedures have been introduced by many authors such as Bahadur [2], Bahadur and Robbins [3], Bechhofer [4], [5], Bechhofer and Blumenthal [6], Bechhofer, Dunnett and Sobel [7], Dunnett [8], Fabian [9], Girshick [10], Gupta and Sobel [11], Mosteller [13] and Paulson [14]~[18] Taylor and David [25], Truax [26]. They have discussed this problem from several different viewpoints such as slippage aspect grouping aspect and ranking aspect. Some of them appealed to sequential multiple decision procedures which were not necessary closed. On the other hand authors such as Armitage [1], Schneiderman [21], Schneiderman and Armitage [22], Sobel and Wald [23], Sobel [24] appealed to restricted or closed sequential procedures which, however, were not necessarily multiple decision procedure.

It was Paulson [19], [20] who presented a class of closed sequential multiple decision procedures for a set of significance level α in $0 < \alpha < 1$ and for a certain configuration of population means for which the probability by which the best population (having the largest means) among several normal populations is selected is larger than the prescribed value $1-\alpha$.

The object of this paper is to generalize the results of Paulson [19], [20] in two directions. In the first place we shall be concerned with a more general class of population distributions, that is, one parameter exponential distributions. In the second place we shall discuss with a configuration of population parameters which are more general than that which Paulson [19], [20] did substantially consider. Such a generalized configuration of population parameters is indeed both subtle and necessary

in our generalized set-up dealing with one parameter exponential distribution.

It is shown in this paper that the essential aspect of our closed sequential statistical procedures (CSSP) can be established in the frame of additive family of sufficient statistics whose notion was introduced by Kitagawa [12]. It is also noted that the restrictions on population distributions are required to establish our results. However these restrictions are so mild that we can easily obtain various examples regarding normal distributions, χ^2 -distributions, Poisson distributions and binomial distributions.

§2. Definitions and introduction.

Definitions. (I). Let Π_i be a population whose probability density function with respect to a common measure μ over the real line is given by

$$(2.1) \quad f(\mathbf{u}_i, \tau_i) d\mu(\mathbf{u}_i) = \exp\{\tau_i \mathbf{u}_i + b(\tau_i) + a(\mathbf{u}_i)\} d\mu(\mathbf{u}_i),$$

defined over $-\infty < \mathbf{u}_i < \infty$ where τ_i is a real parameter whose value is unknown to us, where $b(\tau_i)$ and $a(\mathbf{u}_i)$ are real valued known functions of τ_i and \mathbf{u}_i , respectively, and where $b(\tau_i)$ has the first and the second derivatives $b'(\tau_i)$ and $b''(\tau_i)$ which are continuous and $b'(\tau_i)$ is strictly monotone decreasing with respect to τ_i in a certain finite interval of τ_i ($i=1, 2, \dots, k$). A population is said to be of an exponential type when its probability density function is the form (2.1).

Notation. (I). Let order k values $\tau_1, \tau_2, \dots, \tau_k$ in the ascending order of magnitude, and write in the form $\tau_{(1)} \leq \tau_{(2)} \leq \dots \leq \tau_{(k)}$.

Definition. (II). In what follows in this paper, the best population among k populations of exponential type means the population which has the largest parameter value $\tau_{(k)}$ in case $\tau_{(1)} \leq \tau_{(2)} \leq \dots \leq \tau_{(k)}$ and it means a population which has been selected by certain chance mechanism among s populations whose values of the parameters are equal to $\tau_{(k)}$, when it holds true that $\tau_{(1)} \leq \dots \leq \tau_{(k-s)} < \tau_{(k-s+1)} = \dots = \tau_{(k)}$ for some $s > 1$.

Notation. (II). We denote by $\delta_d(\tau_{(k)}; \lambda_1, \dots, \lambda_{k-1}) = \delta_d(\tau_{(k)}, \lambda)$ where $\lambda = (\lambda_1, \dots, \lambda_{k-1})$, and $\lambda_i \geq 0$, ($i=1, 2, \dots, k-1$), a configuration of population parameters satisfying the conditions $\tau_i = \tau_{(k)} - \Delta - \lambda_i$ for a positive constant value Δ ($i=1, 2, \dots, k-1$).

In what follows we shall be concerned with a certain type of closed sequential statistical procedure whose specification will be given in § 3. A general aspect of any closed sequential statistical procedure may be enunciated as follows:

- (1) There is an assigned positive integer $N \geq 2$.
- (2) There is an assigned sequence of positive integers (m_1, m_2, \dots, m_N) .
- (3) Let us consider k populations and let us draw a sample of size m_1 from each population. Let us construct a rejection region from k

populations which can be determined by means of a preassigned constant of significance level α and observed values of the sample. Then we decide whether each population is rejected or not, by a test defined by the rejection region and by the sample.

(a) In this case, if only one population was not rejected, then we do not draw any more sample and we can and we shall decide that population as the best one.

(b) If more than one populations were left as the candidates for the best one, then we shall proceed to the following second stage.

(ii). In the second stage, let us consider those populations which were not rejected at the first stage enunciated in (i). Let us draw an additional sample of size m_2 from each population and make a pooled sample of size $m_1 + m_2$ for each population. Let us construct a rejection region for populations which can be determined by means of a preassigned constant of significance level α and observed values of the sample. Then we decide whether each population is rejected or not, by a test defined by the rejection region and by the sample.

(a) In this case, if only one population was not rejected, then we do not draw any more sample and we can and we shall decide that population as to be the best one.

(b) If more than one populations were left as the candidates for the best one, then we shall proceed to the following third stage and so on.

(l°). ($3 \leq l \leq N$). We continue these steps in a sequential way, so far as the best population has not been decided and the step number l is smaller than a predetermined integer N .

(($N+1$) $^\circ$). If more than one populations were left as the candidates for the best one at the N -th stage, we shall decide uniquely the best population among those populations which have not been rejected until the N -th stage, by use of certain chance mechanism.

§3. The purpose of this paper.

Definition. (III). We denote by $\mathfrak{F}(T, d)$ a class of configuration $\delta_d(\tau_{(k)}, \lambda)$ of k populations which have the probability density functions of exponential type satisfying the following conditions:

- (1) The parameter $\tau_{(k)}$ belongs to a certain finite interval T .
- (2) The set of $(k-1)$ parameter $\tau_i = \tau_{(k)} - d - \lambda_i$ ($i=1, 2, \dots, k-1$) satisfies the condition that

$$(3. 1) \quad \text{Min}_{1 \leq i \leq k-1} (b'(\tau_i) - b'(\tau_{(k)})) = d(\tau_{(k)}, \lambda) \geq d.$$

Under these definitions and assumptions, the main results of this paper are concerned with the following two purposes (1) and (2):

- (1) To give a family of closed sequential statistical procedure such that, for each of them, the probability by which the best population is selected

is larger than $1-\alpha$ for each assigned α in $0 < \alpha < 1$ and for each set of k populations of exponential type having a configurations $\mathfrak{F}(T, d)$, where the interval T and the positive constant d are arbitrarily preassigned. We shall denote such a class of closed sequential statistical procedures by CSSP $(\alpha; d, T, d)$. (2) To specify a closed sequential statistical procedure which belongs to a class CSSP $(\alpha; d, T, d)$ and which has the minimum number of stages for all configurations belonging to $\mathfrak{F}(T, d)$. However it is hard to establish this object, therefore we shall present several examples in § 7. In what follows, the constants k and d which are enunciated in the previous sections are assumed to be given and fixed.

§4. The enunciation of the closed sequential statistical procedure.

Now our closed sequential procedure, which is denoted by S_θ for each value of θ in the interval $0 < \theta < e_0$, is defined as follows:

(0) Let us define

$$(4.1) \quad N(\theta) = \frac{\log((k-1)/\alpha)}{-\theta h_0(\theta) d}$$

and

$$(4.2) \quad A^{(n)}(\theta) = K(\theta) + n\theta, \quad n=1, 2, \dots, N(\theta),$$

where we put

$$(4.3) \quad K(\theta) = \frac{\log((k-1)/\alpha)}{h_0(\theta) d},$$

and where $h_0(\theta)$ is given as a root of the equation (6.1.19) explained afterwards.

Since the size m_i of sample drawn at the i -th stage is equal to a certain prescribed constant m_i for $i=1, 2, \dots, N(\theta)$, we can and we shall assume that $m_1 = m_2 = \dots = m_{N(\theta)} = 1$ without loss of generality.

(i) In the first stage, let $u_i^{(1)} = U_i^{(1)}$ be a statistic defined by a sample of size one drawn from the population $\Pi_i^{(1)}$, ($i=1, 2, \dots, k$) respectively. Let us arrange the set of k values $U_1^{(1)}, \dots, U_k^{(1)}$ in the ascending order of magnitude, and write them in the form $U_{(1)}^{(1)} \leq \dots \leq U_{(k)}^{(1)}$. If the following relation holds true

$$(4.4) \quad U_{(1)}^{(1)} < U_{(k)}^{(1)} + A^{(1)}(\theta),$$

then we reject the population $\Pi_{(i)}^{(1)}$, in $1 \leq i \leq k-1$. Now our statistical procedure proceeds in either one of two alternative ways: (i)(a) and (i)(b).

(i)(a) In this case, if only one population $\Pi_{(k)}^{(1)}$ was not rejected, then we do not draw any more sample and we can and we shall decide that population $\Pi_{(k)}^{(1)}$ as to be the best one.

(i)(b) If more than one populations were left as the candidates for

the best one, then we shall proceed to the following second stage.

(ii) In the second stage, let us consider those populations which were not rejected at the first stage (i), and let $\{U_{ji}^{(1)}\}$ ($i=1, 2, \dots, k_1$) be a set of k_1 statistics each of which $U_{ji}^{(1)}$ came from the population $\Pi_{ji}^{(1)}$ which was not rejected at the first stage (i), ($i=1, 2, \dots, k_1$). Let us now rewrite $U_{ji}^{(1)}$ and $\Pi_{ji}^{(1)}$ as $W_i^{(2)}$ and $\Pi_i^{(2)}$ ($i=1, 2, \dots, k_1$), respectively. Let $u_i^{(2)}$ be a statistic defined by an additional sample of size one drawn from $\Pi_i^{(2)}$, ($i=1, 2, \dots, k_1$), and put that

$$(4.5) \quad U_i^{(2)} = W_i^{(2)} + u_i^{(2)}, \quad (i=1, 2, \dots, k_1).$$

Let us arrange the set of k_1 values $U_1^{(2)}, \dots, U_{k_1}^{(2)}$ in the ascending order of magnitude, and write them in the form $U_{(1)}^{(2)} \leq \dots \leq U_{(k)}^{(2)}$. If the following relation holds true

$$(4.6) \quad U_{(i)}^{(2)} < U_{(k)}^{(2)} + A^{(2)}(\theta),$$

then we reject the population $\Pi_{(i)}^{(2)}$ with the statistic $U_{(i)}^{(2)}$ ($i=1, 2, \dots, k_1-1$) respectively. Now our statistical procedure proceeds in either one of two alternative ways: (ii) (a) and (ii) (b).

(ii) (a) In this case, if only one population $\Pi_{(k_1)}^{(2)}$ was not rejected, then we do not draw any more sample and we can and we shall decide that population $\Pi_{(k_1)}^{(2)}$ as to be the best one.

(ii) (b) If more than one population were left as the candidates for the best one, then we shall proceed to the following third stage and so on.

(l) ($3 \leq l \leq N(\theta)$). In the l -th stage, let us consider those populations which were not rejected at the $(l-1)$ -th stage and let $\{U_{ji}^{(l-1)}\}$ for $i=1, 2, \dots, k_{l-1}$ be a set of k_{l-1} statistics each of which $U_{ji}^{(l-1)}$ came from the population $\Pi_{ji}^{(l-1)}$ which was not rejected at the $(l-1)$ -th stage, ($i=1, 2, \dots, k_{l-1}$). Let us now rewrite $U_{ji}^{(l-1)}$ and $\Pi_{ji}^{(l-1)}$ as $W_i^{(l)}$ and $\Pi_i^{(l)}$, $i=1, 2, \dots, k_l$ respectively. Let $u_i^{(l)}$ be a statistic defined by an additional sample of size one drawn from $\Pi_i^{(l)}$ ($i=1, 2, \dots, k_{l-1}$) and put that

$$(4.7) \quad U_i^{(l)} = W_i^{(l)} + u_i^{(l)}, \quad (i=1, 2, \dots, k_{l-1}).$$

Let us arrange the set of k_{l-1} values $U_1^{(l)}, \dots, U_{k_{l-1}}^{(l)}$ in the ascending order of magnitude, and write them in the form $U_{(1)}^{(l)} \leq \dots \leq U_{(k_{l-1})}^{(l)}$. If the following relation holds true

$$(4.8) \quad U_i^{(l)} < U_{(k_{l-1})}^{(l)} + A^{(l)}(\theta),$$

then we reject the population $\Pi_i^{(l)}$ with the statistic $U_i^{(l)}$, ($i=1, 2, \dots, k_{l-1}$) respectively. Our statistical procedure proceeds in either one of two alternative way: (l) (a) and (l) (b).

(l) (a) In this case, if only one population $\Pi_{(k_{l-1})}^{(l)}$ was not rejected, then we do not draw any more sample and we can and we shall decide that population $\Pi_{(k_{l-1})}^{(l)}$ as to be the best one.

(l) (b) If more than one populations were left as the candidates for the best one, then we shall proceed to the following stage. We continue

these steps in a sequential ways, so far as the best population has not been decided and until the step number l is smaller than or equal to a predetermined integer $N(\theta)$. If we came to the case of $(N(\theta)^+)$ (b), then we proceed to the following $(N(\theta)+1)$ -th stage.

$((N(\theta)+1)^+)$. In this stage, let us consider those populations which were not rejected at the $N(\theta)$ -th stage and let $\{U_{ji}^{(N(\theta))}\}$ for $i=1, 2, \dots, k_{N(\theta)}$ be a set of $k_{N(\theta)}$ statistics each of which $U_{ji}^{(N(\theta))}$ came from the populations $\Pi_{ji}^{(N(\theta))}$ which was not rejected at the $N(\theta)$ -th stage, ($i=1, 2, \dots, k_{N(\theta)}$). Let us now rewrite $U_{ji}^{(N(\theta))}$ and $\Pi_{ji}^{(N(\theta))}$ as $W_i^{(N(\theta)+1)}$ and $\Pi_i^{(N(\theta)+1)}$ ($i=1, 2, \dots, k_{N(\theta)}$) respectively. Let $U_i^{(N(\theta)+1)}$ be a statistic defined by an additional sample of size one drawn from $\Pi_i^{(N(\theta)+1)}$ ($i=1, 2, \dots, k_{(N(\theta)+1)}$) and put that

$$(4.9) \quad U_i^{(N(\theta)+1)} = W_i^{(N(\theta)+1)} + u_i^{(N(\theta)+1)},$$

$i=1, 2, \dots, k_{N(\theta)}$. Let us arrange the set of $k_{N(\theta)}$ values $U_1^{(N(\theta)+1)}, \dots, U_{k_{N(\theta)}}^{(N(\theta)+1)}$ in the ascending order of magnitude, and write them in the form $U_{(1)}^{(N(\theta)+1)} \leq \dots \leq U_{(k_{N(\theta)})}^{(N(\theta)+1)}$. Then we can and we shall decide the population $\Pi_{(k_{N(\theta)})}^{(N(\theta)+1)}$ as to be the best one.

In this case, when we use the closed sequential procedure S_θ for k populations each of which belongs to the exponential type, we can obtain main theorem which we shall enunciate it in the next section.

§5. The enunciation of theorem.

In order to enunciate a main theorem, let us introduce the following notations:

Let us denote by $P[\Pi_{(k)} S_\theta, \delta_d(\tau_{(k)}, \lambda)]$ the probability by which the best populations of exponential type is rejected by a certain specified closed sequential statistical procedure S_θ under the configuration of population parameters $\delta_d(\tau_{(k)}, \lambda)$.

Then we have the following

Theorem. For a set of constants $\{k, \alpha, \theta, \Delta\}$ which were assigned previously in $0 < \alpha < 1$, $k \geq 2$, $0 < \theta < d$ and $\Delta > 0$, we have

$$(5.1) \quad P[\Pi_{(k)} S_\theta, \delta_d(\tau_{(k)}, \lambda)] \leq \alpha.$$

§6. Proof of the theorem.

6.1. Preliminaries.

In this section let us prove the main theorem in case when $k=2$ and when $k>2$ separately. This is done not only because the case when $k=2$ is important in itself, but also because the probability by which the best population among k populations of exponential type having a configuration of k population parameters $\delta_d(\tau_{(k)}, \lambda)$ is rejected is not larger than the $(k-1)$ times of the probability by which the best population between two populations is rejected under the same parameter configuration, as we shall

varify later.

Before we enter into the proof of the theorem in the case when $k=2$, let us enunciate the following preparations:

First of all let us assume $\tau_2 > \tau_1$ and let us put

$$(6. 1. 1) \quad f(u_{i,j}, \tau_h) := \exp\{\tau_h u_{i,j} + b(\tau_h) + a(u_{i,j})\},$$

for $i, h=1, 2; j=1, 2, \dots$ and put that for each θ in $0 < \theta < d$

$$(6. 1. 2) \quad z_j = \log \left\{ \frac{f(u_{2,j}, \tau_2) f(u_{1,j} + \theta, \tau_1)}{f(u_{2,j}, \tau_1) f(u_{1,j} + \theta, \tau_2)} \right\}, \quad j=1, 2, \dots$$

Lemma 1. (Wald [27], p. 158) *Let z be a random variable such that the following three conditions are fulfilled:*

(A) *The expected value $E\{z\}$ exists and is not equal to 0.*

(B) *There exists a positive ε such that $P[\exp z < 1 - \varepsilon] > 0$ and $P[\exp z > 1 + \varepsilon] > 0$.*

(C) *For any real value h the expected value $E\{\exp(hz)\} = g(h)$ exists.*

Then there exists one and only one real value $h_0 \neq 0$ such that

$$(6. 1. 3) \quad E\{\exp(h_0 z)\} = 1.$$

Lemma 2. *Let $\{z_j\}$ be a set of independent random variables for which (6. 1. 1) and (6. 1. 2) are defined. Then each random variable z_j in a set of $\{z_j\}$ satisfies the three conditions (A), (B) and (C) enunciated in Lemma 1.*

Proof: Ad(A). For each z_j we have

$$(6. 1. 4) \quad \begin{aligned} E\{z_j\} &= (\tau_2 - \tau_1) E\{u_{2,j} - u_{1,j} - \theta\} \\ &= (\tau_2 - \tau_1) \{b'(\tau_1) - b'(\tau_2) - \theta\} \\ &\geq \Delta(d - \theta) > 0, \end{aligned}$$

under the configuration $\delta_d(\tau_2, \lambda_1)$, where Δ and d are two specified positive constant values given in §3.

Ad(B). For any assigned constant value of θ in $0 < \theta < d$ and for any assigned constant value of ε in $0 < \varepsilon < 1$, if we have that the following probability

$$(6. 1. 5) \quad \begin{aligned} &P[\exp z_j < 1 - \varepsilon \mid \delta_d(\tau_2, \lambda_1)] \\ &= P[u_{2,j} < u_{1,j} + \theta + (\tau_2 - \tau_1)^{-1} \log(1 - \varepsilon) \mid \delta_d(\tau_2, \lambda_1)] \end{aligned}$$

is equal to 0, then the following equation holds true

$$(6. 1. 6) \quad P[u_{2,j} \geq u_{1,j} + \theta + (\tau_2 - \tau_1)^{-1} \log(1 - \varepsilon) \mid \delta_d(\tau_2, \lambda_1)] = 1,$$

for any constant values θ and ε in $0 < \theta < d$ and in $0 < \varepsilon < 1$. This contradicts to the definition of the measure μ , since the random variables $u_{1,j}$ and $u_{2,j}$ are mutually independent.

Similarly we can verify the second part of (B).

Ad(C). For any real value of h we have

$$\begin{aligned}
(6.1.7) \quad g(h) &:= E\{\exp\{h(\tau_2 - \tau_1)(u_{2j} - u_{1j} - \theta)\}\} \\
&= \int_{-\infty}^{\infty} \exp\{(\tau_2 - \tau_1)h + \tau_2\{u_{2,j} + b(\tau_2) + a(u_{2,j})\}\} d\mu(u_{2,j}) \\
&\quad \times \int_{-\infty}^{\infty} \exp\{\{\tau_1 - \tau_2\}h + \tau_1\{u_{1,j} + b(\tau_1) + a(u_{1,j})\}\} d\mu(u_{1,j}) \\
&\quad \times \exp\{(\tau_1 - \tau_2)h\theta\} \\
&= \exp\{b(\tau_1) + b(\tau_2) + (\tau_1 - \tau_2)h\theta - b((\tau_2 - \tau_1)h + \tau_2) \\
&\quad - b((\tau_1 - \tau_2)h + \tau_1)\}.
\end{aligned}$$

The immediate consequence of Lemma 2 is the following

Lemma 3. *Let $\{z_j\}$ be a set of independent random variables for which (6.1.1) and (6.1.2) are defined. Then there exists one and only one real value $h_0 \neq 0$ such that*

$$(6.1.8) \quad E\{\exp\{h_0 z_j\}\} = 1, \quad (j=1, 2, \dots).$$

Lemma 4. *The value h_0 in (6.1.8) is negative.*

Proof: In virtue of (6.1.4), the proof is immediate from the proof of Lemma 1 (Wald [27], p. 159).

The immediate consequence of Lemma 3 and Lemma 4 is the following

Lemma 5. *Let $B(\theta)$ be any assigned positive function defined over the interval $0 < \theta < d$. Let us terminate the sampling at the n -th stage, if*

$$(6.1.9) \quad B(\theta) < \exp\{Z^{(n)}\}, \quad (m=1, 2, \dots, n-1),$$

and

$$(6.1.10) \quad \exp\{Z^{(n)}\} \leq B(\theta),$$

where $z_j, j=1, 2, \dots, n$ are defined by (6.1.1) and (6.1.2), and where we have put $Z^{(n)} = \sum_{j=1}^n z_j$.

Then we have

(1) *For each θ in $0 < \theta < d$ the value of $h_0(\theta)$ is determined so that $h_0(\theta) < 0$ and that*

$$(6.1.11) \quad E\{\exp\{Z^{(n)} h_0(\theta)\}\} = 1.$$

(2) *For every $n \geq 1$ we have*

$$(6.1.12) \quad B(\theta)^{h_0(\theta)} \int \dots \int \prod_{i=1}^2 \prod_{j=1}^n f(u_{i,j}, \tau_j) d\mu(u_{\cdot,j}) \leq 1.$$

$\exp\{Z^{(n)}\} \leq B(\theta)$

Proof: Ad (1) Since $\{z_j\}$ is a set of independent and identically distributed random variables, we have

$$(6.1.13) \quad E\{\exp\{Z^{(n)} h_0(\theta)\}\} = \prod_{j=1}^n E\{\exp\{z_j h_0(\theta)\}\} = 1.$$

Ad(2). In virtue of (1), we have

$$\begin{aligned}
 (6. 1. 14) \quad 1 &= E\{ \exp\{Z^{(n)}h_0(\theta)\} \} \\
 &= \int \cdots \int \exp\{Z^{(n)}h_0(\theta)\} \prod_{i=1}^2 \prod_{j=1}^n f(u_{i,j}, \tau_j) d\mu(u_{i,j}) \\
 &\quad \exp\{Z^{(n)} \leq B(\theta)\} \\
 &\geq B(\theta)^{h_0(\theta)} \int \cdots \int \prod_{i=1}^2 \prod_{j=1}^n f(u_{i,j}, \tau_j) d\mu(u_{i,j}). \\
 &\quad \exp\{Z^{(n)} \leq B(\theta)\}
 \end{aligned}$$

Lemma 6. (Paulson [19], p.176) *If $\{Y_j\}$ is a sequence of independent and identically distributed random variables with a negative expectation and $b > 0$, then we have*

$$(6. 1. 15) \quad P\left[\sum_{j=1}^n Y_j > b \text{ for some } n < \infty \right] \leq \exp\{-h_0 b\},$$

where h_0 is the non-zero root of $E\{\exp\{hY\}\} = 1$.

Lemma 7. *Let us put $Z^{(n)} = \sum_{j=1}^n z_j$ for $n = 1, 2, \dots$, where $\{z_j\}$ are defined by (6. 1. 1) and (6. 1. 2). Let $B(\theta)$ be any assigned positive, however smaller than one, function defined over the interval $0 < \theta < d$. Let $h_0(\theta)$ be a value such that $h_0(\theta) < 0$ and that (6. 1. 11) is satisfied. Then we have*

$$(6. 1. 16) \quad P[Z^{(n)} \leq \log B(\theta) \text{ for some } n < \infty] \leq B(\theta)^{-h_0(\theta)}$$

Proof: In Lemma 6, let us put that $b = -\log B(\theta)$ and $Y_j = -z_j$ ($j = 1, 2, \dots$). By use of (6. 1. 4), the relation (6. 1. 16) follows immediately from the relation (6. 1. 15).

Lemma 8. *For each assigned θ in $0 < \theta < d$ and for any assigned α in $0 < \alpha < 1$, let us assume that the following relation*

$$(6. 1. 17) \quad -\log B(\theta) \geq h_0(\theta)^{-1} \log \alpha$$

hold true. Then we have

$$(6. 1. 18) \quad P[Z^{(n)} \leq \log B(\theta) \text{ for some } n < \infty] \leq \alpha.$$

Proof: The proof of this lemma follows from Lemma 7 immediately.

Lemma 9. *For each assigned θ in $0 < \theta < d$, the root $h_0(\theta)$ defined by (6. 1. 11) is given by the root of the following functional equation*

$$(6. 1. 19) \quad b(\tau_1) + b(\tau_2) = b(\tau_1^*) + b(\tau_2^*) + \theta h_0(\theta) (\tau_2 - \tau_1),$$

where

$$(6. 1. 20) \quad \tau_1^* = \tau_1 + (\tau_1 - \tau_2) h_0(\theta),$$

$$(6. 1. 21) \quad \tau_2^* = \tau_2 + (\tau_2 - \tau_1) h_0(\theta).$$

Proof: For each assigned θ in $0 < \theta < d$, the proof follows immediately

from the substitution of $h_0(\theta)$ to h in (6. 1. 7).

6. 2. The proof of the theorem in the case when $k=2$.

First of all, in order to prove the theorem when $k=2$ let us introduce the following notations for the sake of convenience:

For each l ($1 \leq l \leq N(\theta) + 1$), let us denote by $D_l^{(l)}((1, 2); 1)$ the stochastic event by which both population Π_1 and Π_2 are not rejected before the l -th stage and which the population Π_2 is rejected at the l -th stage by our statistical procedure S_θ . Let us denote by $D_{N(\theta)+1}((1, 2); 1)$ the stochastic event by which the population Π_2 is rejected at a certain stage by our statistical procedure S_θ .

In this connection let us denote by $P[\Pi_{(2)} | S_\theta, \delta_d(\tau_{(2)}, \lambda)]$ the probability by which the population Π_2 is rejected by our statistical procedure S_θ , under the configuration of population parameters $\delta_d(\tau_{(2)}, \lambda)$. Then we have the following

Theorem. *We have for each assigned α in $0 < \alpha < 1$*

$$(6. 2. 1) \quad P[\Pi_{(2)} | S_\theta, \delta_d(\tau_{(2)}, \lambda)] \leq \alpha.$$

Proof: Firstly let us consider the case when $\tau_2 = \tau_{(2)}$. Then we have

$$(6. 2. 2) \quad \begin{aligned} & P[\Pi_{(2)} | S_\theta, \delta_d(\tau_{(2)}, \lambda)] \\ &= P[D_{N(\theta)+1}((1, 2); 1) | S_\theta, \delta_d(\tau_{(2)}, \lambda)] \\ &= P[D_n^{(n)}((1, 2); 1) \text{ for some } n \ (1 \leq n \leq N(\theta) + 1) | S_\theta, \delta_d(\tau_{(2)}, \lambda)] \\ &= P\left[\sum_{i=1}^n z_i < (d + \lambda)K(\theta) \text{ for some } n < \infty \mid S_\theta, \delta_d(\tau_{(2)}, \lambda)\right], \end{aligned}$$

where we have put

$$(6. 2. 3) \quad z_i = (\tau_2 - \tau_1)(u_{2,i} - u_{1,i} - \theta), \quad (i=1, 2, \dots).$$

For each assigned θ in $0 < \theta < d$, let us consider the root $h_0(\theta)$ of the following equation

$$(6. 2. 4) \quad E[\exp(z h_0(\theta))] = 1.$$

Since we have

$$(6. 2. 5) \quad \begin{aligned} E[z] &= (\tau_2 - \tau_1)(b'(\tau_1) - b'(\tau_2) - \theta) \\ &\geq d(d - \theta) > 0, \end{aligned}$$

there exists one and only one negative root $h_0(\theta)$ of (6. 2. 4) in virtue of Lemma 4. In virtue of (6. 1. 18) the probability in the right hand side of (6. 2. 2) does not exceed the following

$$(6. 2. 6) \quad \begin{aligned} & \exp\{-h_0(\theta)(d + \lambda)K(\theta)\} \\ &= \exp\{-(d + \lambda)/d \log(1/\alpha)\} \leq \alpha, \end{aligned}$$

where $h_0(\theta)$ is calculated from (6. 1. 19) ~ (6. 1. 21).

In case when $\tau_1 = \tau_{(2)}$, we can also verify the theorem by the similar discussions to that in case $\tau_2 = \tau_{(2)}$.

6. 3. The proof of the theorem in case when $k > 2$.

Let us denote by $\delta_d(\tau_k, \lambda)$ the parameter configuration $\tau_{(k)} = \tau_k = \tau_j + \Delta + \lambda_j$, $\lambda_j \geq 0$ for $j = 1, 2, \dots, k-1$. Paulson [19] has used the following result under our notations.

Lemma 10. (Paulson [19], p. 176) *we have*

$$(6. 3. 1) \quad P[\Pi_k | S_\theta, \delta_d(\tau_k, \lambda)] \leq \sum_{j=1}^{k-1} P[U_k^{(n)} - U_j^{(n)} < A^{(n)}(\theta) \text{ for some } n (1 \leq n \leq N(\theta) + 1) | S_\theta, \delta_d(\tau_k, \lambda)].$$

Lemma 11. *We have*

$$(6. 3. 2) \quad P[\Pi_{(k)} | S_\theta, \delta_d(\tau_{(k)}, \lambda)] = P[\Pi_k | S_\theta, \delta_d(\tau_k, \lambda)].$$

Proof: Since our closed sequential statistical procedure S_θ defined by §4 is symmetric with regard to $\tau_{(k)}$, the proof follows immediately.

Hence we have the following main theorem enunciated in §5.

Theorem. *For a set of constants $\{k, \alpha, \theta, \Delta\}$ which were assigned previously in $0 < \alpha < 1$, $k \geq 2$, $0 < \theta < d$ and $\Delta > 0$, we have*

$$(6. 3. 3) \quad P[\Pi_{(k)} | S_\theta, \delta_d(\tau_{(k)}, \lambda)] \leq \alpha.$$

Proof: The proof is immediate from Lemma 10, Lemma 11 and (6. 2. 1).

Lemma 12. *For any assigned integers j_i in $1 \leq j_i \leq k-1$, $i = 1, 2$; $j_1 \neq j_2$, we have*

$$(6. 3. 4) \quad P[U_k^{(n)} - U_{j_1}^{(n)} < A^{(n)}(\theta) \text{ for some } n (1 \leq n \leq N(\theta) + 1) | S_\theta, \delta_d(\tau_k, \mathbf{0})] = P[U_k^{(n)} - U_{j_2}^{(n)} < A^{(n)}(\theta) \text{ for some } n (1 \leq n \leq N(\theta) + 1) | S_\theta, \delta_d(\tau_k, \mathbf{0})]$$

Proof: For any assigned numbers n , j_1 and j_2 in $1 \leq n \leq N(\theta) + 1$ and in $1 \leq j_i \leq k-1$, $i = 1, 2$; $j_1 \neq j_2$, the random variables $U_{j_1}^{(n)}$ and $U_{j_2}^{(n)}$ are identically distributed under the parameter configuration $\delta_d(\tau_k, \mathbf{0})$. Hence the proof is immediate.

Corollary. *We have*

$$(6. 3. 5) \quad P[\Pi_{(k)} | S_\theta, \delta_d(\tau_{(k)}, \mathbf{0})] \leq (k-1) P[\Pi_{(2)} | S_\theta, \delta_d(\tau_{(2)}, \mathbf{0})].$$

Proof: The proof is immediate from Lemma 11 and Lemma 12.

§7. Examples.

Let us give several examples concerning with the specification of the closed sequential statistical procedure which has the minimum number of stages with regard to θ in $0 < \theta < d$, where d is a preassigned constant

which was defined by §3.

Example 1. (Mean of normal distribution). Let $\{u_{i,j}\}$, ($i=1, 2, \dots, k$; $j=1, 2, \dots$) be k sequences of independent statistics having their probability density function of (2. 1), provided that

$$(7. 1) \quad \tau_i = \mu_i, \quad b(\tau_i) = -\tau_i^2/2, \quad (i=1, 2, \dots, k).$$

In this case let us determine a value of $N(\theta)$, defined by (4. 1), with regard to θ in $0 < \theta < \Delta$, where Δ is a preassigned positive constant defined by §2.

In virtue of (6. 1. 19) we have

$$(7. 2) \quad h_0(\theta) = -1 + \theta/\Delta,$$

and we have

$$(7. 3) \quad \text{Min}_{0 < \theta < \Delta} \{ \theta h_0(\theta) \} = \Delta/4$$

therefore we have

$$(7. 4) \quad \text{Min}_{0 < \theta < \Delta} N(\theta) = 4 \log((k-1)/\alpha) / \Delta^2.$$

Example 2. (Variance of normal distribution). Let $\{u_{i,j}\}$, ($i=1, 2, \dots, k$; $j=1, 2, \dots$) be k sequences of independent statistics having their probability density function (2. 1), provided that

$$(7. 5) \quad \tau_i = -(2\sigma_i^2)^{-1}, \quad (\sigma_i > 0), \quad b(\tau_i) = 2 \log(-\tau_i),$$

($i=1, 2, \dots, k$). In virtue of (8. 1. 19) we have

$$(7. 6) \quad -\theta h_0(\theta) = \log \{ (\tau_{(k)} + \Delta h_0(\theta)) (\tau_{(k)} - \Delta h_0(\theta)) / ((\tau_{(k)} - \Delta) \tau_{(k)}) \} / (2\Delta).$$

Now let us put that

$$(7. 7) \quad \theta_{0,2} = \Delta^{-1} \log \{ (2\tau_{(k)} - \Delta)^2 / (4\tau_{(k)}(\tau_{(k)} - \Delta)) \},$$

$$(7. 8) \quad d_2 = (1/(\tau_{(k)} - \Delta) - 1/\tau_{(k)})/2.$$

Then we have

$$(7. 9) \quad \text{Min}_{0 < \theta < d_2} N(\theta) = 2 \log((k-1)/\alpha) / (\theta_{0,2} \Delta).$$

Example 3. (Scale parameter of exponential distribution). Let $\{u_{i,j}\}$, ($i=1, 2, \dots, k$; $j=1, 2, \dots$) be k sequences of independent statistics having their probability density function of (2. 1), provided that

$$(7. 10) \quad \tau_i = -\lambda_i^{-1}, \quad (\lambda_i > 0), \quad b(\tau_i) = \log(-\tau_i), \quad (i=1, 2, \dots, k).$$

By means of the similar method used in Example 2, we have the following minimum value of $N(\theta)$ with regard to θ in $0 < \theta < d_3$, where we put

$$(7. 11) \quad d_3 = 1/(\tau_{(k)} - \Delta) - 1/\tau_{(k)}.$$

$$(7.12) \quad \underset{0 < \theta < d_1}{\text{Min}} N(\theta) = 2 \log((k-1)/\alpha) / (\theta_{0,3} d).$$

where we have put

$$(7.13) \quad \theta_{0,3} = \frac{2}{d} \log \left\{ \frac{(2\tau_{(k)} - d)^2}{4\tau_{(k)}(\tau_{(k)} - d)} \right\}.$$

Example 4. (Poisson distribution). Let $\{u_{i,j}\}$, $i=1, 2, \dots, k$; $j=1, 2, \dots$ be k sequences of independent statistics having their probability density function of (2.1), provided that

$$(7.14) \quad \tau_i = \log \lambda_i, \quad (\lambda_i > 0), \quad b(\tau_i) = -\lambda_i, \quad (i=1, 2, \dots, k).$$

Let us put that

$$(7.15) \quad \theta_{0,4} = 2[\exp(\tau_{(k)}/2) - \exp\{(\tau_{(k)} - d)/2\}]^2 / d,$$

$$(7.16) \quad d_4 = (1 - \exp\{-d\}) \exp\{\tau_{(k)}\}.$$

By means of the similar method used in the previous examples, we have the following

$$(7.17) \quad \underset{0 < \theta < d_4}{\text{Min}} N(\theta) = 2 \log((k-1)/\alpha) / (\theta_{0,4} d).$$

Example 5. (Binomial distribution). Let $\{u_{i,j}\}$, $(i=1, 2, \dots, k$; $j=1, 2, \dots)$ be k sequences of independent statistics having their probability density function of (2.1), provided that

$$(7.18) \quad \tau_i = \log(p_i/(1-p_i)), \quad (0 < p_i < 1), \quad b(\tau_i) = \log(1-p_i), \quad (i=1, 2, \dots, k).$$

By means of the similar method used in the previous examples, we have the following. Let us put that

$$(7.19) \quad \theta_{0,5} = \frac{2}{d} \log \left\{ \frac{(1 + \exp\{\tau_{(k)}\})(1 + \exp\{\tau_{(k)} - d\})}{(1 + \exp(\tau_{(k)} - d/2))^2} \right\}$$

$$(7.20) \quad d_5 = (1 + \exp\{\tau_{(k)}\})^{-2} \exp\{\tau_{(k)}\} - (1 + \exp\{\tau_{(k)} - d\})^{-2} \exp\{\tau_{(k)} - d\},$$

then we have

$$(7.21) \quad \underset{0 < \theta < d_5}{\text{Min}} N(\theta) = 2 \log((k-1)/\alpha) / (\theta_{0,5} d).$$

Acknowledgement.

The author would like to express his hearty thanks to Professor T. Kitagawa of Mathematical institute, Kyushu University for his encouragements and criticisms during this paper was prepared.

References.

- [1] Armitage, P.: *Restricted sequential procedures*, *Biometrika*, **44** (1957), 9-26.
 [2] Bahadur, E.: *On c problem in the theory of k populations*, *Ann. Math. Stat.*, **21** (1950), 362-

- 375.
- [3] Bahadur, E. and Robbins, H.: *The problem of the greater mean*, Ann. Math. Stat., **21** (1950), 469-487.
 - [4] Bechhofer, R. E.: *A single sample multiple decision procedure for ranking means of normal populations with known variance*, Ann. Math. Stat., **25** (1954), 16-39.
 - [5] Bechhofer, R. E.: *A sequential multiple-decision procedure for selecting the best one of several normal populations with a common unknown variance, and its use with various experimental designs*, Biometrics, **14** (1958), 408-429.
 - [6] Bechhofer, R. E. and Blumenthal, S.: *A sequential multiple-decision procedure for selecting the best one of several normal populations with a common unknown variance, II; Monte Carlo sampling results and new computing formula*, Biometrika, **18** (1962), 52-67.
 - [7] Bechhofer, R. E., Dunnett, C. W. and Sobel, M.: *A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variances*, Biometrika, **41** (1954), 170-176.
 - [8] Dunnett, C. W.: *A multiple comparison procedure for comparing several treatments with a control*, Jour. Amer. Stat. Assn., **50** (1955), 1096-1121.
 - [9] Fabian, V.: *On multiple decision methods for ranking population means*, Ann. Math. Stat., **33** (1962), 248-254.
 - [10] Girshick, M. A.: *Contributions to the theory of sequential analysis*, Ann. Math. Stat., **17** (1946), 123-143.
 - [11] Gupta, S. S. and Sobel, M.: *On selecting a subset containing the population with the smallest variance*, Biometrika, **49** (1962), 495-507.
 - [12] Kitagawa, T.: *Successive process of statistical inference associated with an additive family of sufficient statistics*, Bull. Math. Stat., **7** (1957), 92-112.
 - [13] Mosteller, F.: *A k-sample slippage test for an extreme population*, Ann. Math. Stat., **19** (1948), 58-65.
 - [14] Paulson, E.: *A Multiple decision procedure for certain problem in the analysis of variance*, Ann. Math. Stat., **20** (1949), 95-98.
 - [15] Paulson, E.: *On the comparison of several experimental categories with a control*, Ann. Math. Stat., **23** (1952), 239-246.
 - [16] Paulson, E.: *An optimum solution to the k sample slippage problem for the normal distribution*, Ann. Math. Stat., **23** (1952), 610-616.
 - [17] Paulson, E.: *A sequential procedure for comparing several experimental categories with a standard or control*, Ann. Math. Stat., **33** (1962), 438-443.
 - [18] Paulson, E.: *A sequential procedure for choosing one of k hypotheses concerning the unknown mean of a normal distribution*, Ann. Math. Stat., **34** (1963), 549-554.
 - [19] Paulson, E.: *A sequential procedure for selecting the population with the largest means from k normal population* Ann. Math. Stat., **35** (1964), 174-180.
 - [20] Paulson, E.: *Sequential estimation and closed sequential decision procedures*, Ann. Math. Stat., **35** (1964), 1048-1058.
 - [21] Schneiderman, M. A.: *Closed sequential t-test*, Biometrika, **49** (1962), 359-366.
 - [22] Schneiderman, M. A. and Armitage, P.: *A family of closed sequential procedures*, Biometrika, **49** (1962), 41-56.
 - [23] Sobel, M. and Wald, A.: *A sequential decision procedure for choosing one of three hypotheses concerning with unknown mean of a normal distribution*, Ann. Math. Stat., **20** (1949), 502-522.
 - [24] Sobel, M.: *Sequential procedures for selecting the best exponential population*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Univ. of California Press, Berkeley, Vol. 5 (1956) 99-110.
 - [25] Taylor, R. J. and David, H. A.: *A multi-stage procedure for the selection of the best of several populations*, Jour. Amer. Stat. Assn., **57** (1962), 785-795.
 - [26] Truax, D. R.: *An optimum slippage test of the variances of k normal populations*, Ann. Math. Stat., **28** (1953), 669-674.
 - [27] Wald, A.: *Sequential analysis*, Wiley, New York (1947).