OPTIMUM DESIGNS FOR SELECTING ONE OF TWO TREATMENTS, FIXED SAMPLE SIZE PLAN AND SEQUENTIAL PLAN IN A POPULATION WITH ONE PARAMETER EXPONENTIAL DISTRIBUTION

Sugimura, Masahiko
Kumamoto Women's University

OPTIMUM DESIGNS FOR SELECTING ONE OF TWO TREATMENTS, FIXED SAMPLE SIZE PLAN AND SEQUENTIAL PLAN IN A POPULATION WITH ONE PARAMETER EXPONENTIAL DISTRIBUTION

By

Masahiko SUGIMURA
(Received Oct. 31, 1967)

Introduction

The problem to determine the optimum statistical procedure, in some specified sense, for choosing between two populations in the light of samples drawn from them, is very important in practical situation.

Let us now enunciate a formulation of this problem in a more concrete form.

In the first place, we shall treat the situation in which there are two treatments (denoted by A and B) to be performed on each of a large number of individuals (say N), and we shall assume that the effect of each treatment on each individual can be expressed in terms of one real number, and that it is distributed in accordance with a certain population distribution.

Let x_A be a treatment effect of A performed on an individual randomly drawn from the population associated with the treatment A, and let us assume that x_A is distributed with a distribution function $F_A(x; \theta_A)$ under our circumstance, where θ_A is a parameter in the population distribution. While let x_B be a treatment effect of B performed on an individual randomly drawn from the other population associated with the treatment B, and let us assume that x_B is distributed with a distribution function $F_B(x; \theta_B)$ under our circumstance, where θ_B is a parameter in the population distribution.

Under this general circumstance we shall be concerned with two types of the problem.

In the problem of Type I, we set up the following assumptions concerning $F_A(x; \theta_A)$ and $F_B(x; \theta_B)$:

(a) The types of the two distribution functions are the same and known to us.

(b) The true values of θ_A and θ_B are both unknown to us.

(c) Let us denote by $d(\theta_A, \theta_B)$ an assigned function of θ_A and θ_B expressing a discrepancy between θ_A and θ_B. There exists an *a priori* distribution for the parameter $\delta = d(\theta_A, \theta_B)$ and it is a continuous type and known to us.

On the other hand, in the problem of Type II, we set up the following assumptions concerning $F_A(x; \theta_A)$ and $F_B(x; \theta_B)$:

1

(a) The types of the two distribution functions are the same and known to us.

(b) The true value of θ_B is known to us, while that of θ_A unknown.

(c) Let us denote by $d(\theta_A, \theta_B)$ an assigned function of θ_A and θ_B expressing a discrepany between θ_A and θ_B. There exists an *a priori* distribution for the parameter $\delta = d(\theta_A, \theta_B)$ and it is a continuous type and known to us.

Now there is another different aspect of our approaches. Indeed there are two different sampling plans, as we shall show.

(i) In **the fixed sample size plan**, we shall consider a fixed number of individuals randomly drawn from the population of individuals.

(ii) In **the sequential plan**, the number of individuals drawn will not be fixed but a sequential statistical procedure will be performed.

Now the combination of two aspects just enunciated will give us at least formally possible four cases:

(1) **The first case** in which the problem of Type I will be carried out under the fixed sample size plan.

(2) **The second case** in which the problem of Type II will be carried out under the fixed sample size plan.

(3) **The third case** in which the problem of Type I will be carried out under the sequential plan.

(4) **The fourth case** in which the problem of Type II will be carried out under the sequential plan.

In this paper, we want to investigate our problem in each of the above-mentioned four cases according to the following common scheme.

(1°) Firstly, a trial will be performed according to the following method peculiar to each of the four cases:

(1) In the first case, a trial will be performed on $2n$ randomly drawn from N individuals, n on each treatment.

(2) In the second case, a trial will be performed for n randomly drawn from N individuals on the treatment A.

(3) In the third case, a trial will be performed successively on each pair randomly drawn from N individuals, one member of the pair on the treatment A and the other on the treatment B.

(4) In the fourth case, a trial will be performed successively for each randomly drawn from N individuals on the treatment A.

(2°) Secondly, after the trial, if either of the two treatments A and B is decided to be *superior* to the other according to a certain statistical procedure in view of the observations supplied by the samples obtained in our trial, then the treatment will be performed on all of the remaining individuals.

It is noted that we shall use the word "*superior*" or "*inferior*" throughout this paper on the basis of the following definition and that the decision that one treatment is superior to the other will be performed according to a certain statistical procedure.

Definition 1: *Let us assume that there exists an a priori defined interval denoted by* $\mathbf{I} = (\varDelta_1, \varDelta_2)$ *of* δ, *which is not necessarily bounded. Let us also divide the interval* $\mathbf{I}$ *into two sub-intervals* $\varOmega_0 = (\varDelta_1, \delta^*)$ *and* $\varOmega_1 = (\delta^*, \varDelta_2)$ *by a specified real value* δ^* *of* δ,

where $J_1 < \delta^* < J_2$. Now, when the true value of the parameter is δ, the treatment A is called to be inferior to the treatment B (, that is, the treatment B is called to be superior to the treatment A) if $\delta \in \Omega_0$ and conversely the treatment A is called to be superior to the treatment B (, that is, the treatment B is called to be inferior to the treatment A) if $\delta \in \Omega_1$.

In considering such a scheme mentioned at (1°) and (2°), we shall proceed to choose the optimum statistical procedure according to a certain criterion of choice among some specified class of procedures.

Under these specified formulations our main interests are now reduced essentially to set up the criterion of choice. Indeed we shall give our criterion from the point of view of decision theory. For the purpose of this decision theoretic approach, let us now define *the loss* associated with our statistical procedure as follows.

Definition 2: *When the true value of the parameter is δ, the loss per one individual inccured from performing A is $C(\delta^* - \delta)$ if $\delta \in \Omega_0$ and it is zero if $\delta \in \Omega_1$, while the loss per one individual inccured from performing B is $C(\delta - \delta^*)$ if $\delta \in \Omega_1$ and it is zero if $\delta \in \Omega_0$, where C is a specified proportionality factor.*

On the basis of this loss for N individuals, we shall now consider the average risk over the *a priori* distribution of the parameter δ which we have already assumed in this paper as to exist and known to the experimenter.

(We shall often call the over-all expected loss for the average risk in this paper.)

We can now express the formulation of our problem in a concrete form as follows.

A formulation of our statistical procedures in this paper is to determine the sample size in the trial and the choice of decision for the fixed sample size plan, while it is to determine the stopping rule and the choice of decision for the sequential plan. In either of these two plans, our criterion is based upon the criterion to minimize the average risk over the *a priori* distribution of the parameter δ.

We have proposed the loss approach to our problem. On the other hand, let us now propose another approach which differs from the loss approach in location of a base line, that is, *the net gain approach*. For the purpose of this approach, we shall give a definition of the net gain associated with our statistical procedure as follows.

Definition 3: *When the true value of the parameter is δ, the gain per one individual obtained from performing A is $G(\delta - \delta^*)$ and the gain per one individual obtained from performing B is $G(\delta^* - \delta)$, where G is a specified proportionality factor.* (It should be noted that $G(\delta - \delta^*)$ is positive if $\delta \in \Omega_1$ and negative if $\delta \in \Omega_0$ and $G(\delta^* - \delta)$ is negative if $\delta \in \Omega_1$ and positive if $\delta \in \Omega_0$.

Thus we can construct the expected net gain for all N individuals on the basis of this gain. In this gain approach, we shall adopt the criterion to maximize the over-all expected net gain for all N individuals, which is obtained by integrating the expected net gain over the *a priori* distribution for δ, instead of the criterion to minimize the average risk. However, on account of our stand point of considering the over-all expected net gain or the average risk over the *a priori* distribution for δ, the maximization of the over-all expected net gain is equivalent to the minimization of the average risk. Therefore, we shall be concerned with one or both of two approaches in each section as occasion demands.

The above-mentioned attempts have been so far performed by Colton [1], [2] for normal distributions in the first and the third cases, and by the author [4]~[12] for normal distributions and binomial distributions in four cases.

The object of this paper is to generalize these results in the following direction. We shall be concerned with a more general class of population distribution, that is, one parameter exponential distribution. Explaining in detail, we formulate our problem in the frame of the independent sufficient statistics admitting *additivity* whose notion was introduced by Kitagawa [3]. In chapter 1, we shall formulate our problem under the fixed sample size plan, and in chapter 2, we shall formulate our problem under the sequential plan. In the final section of each chapter, as its special examples, we shall formulate our problem in the cases of normal, binomial, Poisson, and Gamma distributions.

Assumption

Let us consider a population Π admitting the sufficient statistic $u_m = \sum_{i=1}^{m} v_i$ ($m = 1, 2, 3, \cdots$) for a parameter δ, where $\{v_i\}$ indicates a sequence of independent random variables v_i's drawn from the population Π and each v_i has the probability density function with respect to a common measure μ over the real line R such that

(0.1) $$f(v_i ; \delta)d\mu(v_i) = exp \{-\delta v_i + b(\delta) + a_1(v_i)\} d\mu(v_i) ,$$

where δ is the parameter whose value is unknown to us, where $b(\delta)$ and $a_1(v_i)$ are real valued known functions of δ and v_i respectively, and where $b(\delta)$ has the first and second derivatives which are continuous in an *a priori* preassigned interval $\mathbf{I} = (\varDelta_1, \varDelta_2)$ and $b'(\delta)$ is strictly monotone decreasing in $\mathbf{I}$.

In the first case or the third case, in which the problem of Type I will be performed under the fixed sample size plan or the sequential plan, we hereupon consider the statistic u_m which indicates a discrepancy between the observations from performing half and half the treatments A and B on $2n$ individuals drawn from N individuals.

In the second case or the fourth case, in which the problem of Type II will be performed under the fixed sample size plan or the sequential plan, we consider the statistic u_m which indicates a discrepancy between the observation from performing treatment A on n individuals drawn from N individuals and the known value of θ_B.

We assume that any pair $\{u_m, v_{m+1}\}$ of independent sufficient statistics composed from v_i ($i = 1, 2, 3, \cdots$) with the probability density function of (0.1) admits the *additivity* proposed by Kitagawa [3], where $m = 1, 2, 3, \cdots$.

The necessary and sufficient condition that the additivity be satisfied concerning independent sufficient statistics was given by Kitagawa [3] as follows.

The necessary and sufficient condition that $u_{m+1} = u_m + v_{m+1}$ be a statistic having the probability density function with respect to the measure μ over the real line R such that

(0.2) $$f_{m+1}(u_{m+1} ; \delta)d\mu(u_{m+1}) = exp \{-\delta u_{m+1} + (m+1)b(\delta) + a_{m+1}(u_{m+1})\} d\mu(u_{m+1})$$

for any positive integer m is that

(0.3)
$$exp \{a_{m+1}(u)\} = \int_R exp \{a_m(u-v)\} \cdot exp \{a_1(v)\} d\mu(v)$$

for any positive integer m.

Therefore, we may and will assume that each statistic u_m $(m = 1, 2, 3, \cdots)$ has the probability density function with respect to the measure μ over the real line R such that

(0.4)
$$f_m(u_m; \delta) d\mu(u_m) = exp \{-\delta u_m + m \cdot b(\delta) + a_m(u_m)\} d\mu(u_m),$$

where $a_m(u_m)$ is a real valued known function of u_m for each m $(m = 1, 2, 3, \cdots)$.

CHAPTER 1

FIXED SAMPLE SIZE PLAN IN A POPULATION WITH ONE PARAMETER EXPONENTIAL DISTRIBUTION

§1.1. Procedure.

The First Case: Perform a trial on $2n$ drawn from N individuals, each treatment on n. Compute the observed value of $u_n = \sum_{i=1}^{n} v_i$.

Procedure: If $u_n > K_2$, use treatment B on the remaining $N-2n$;

If $u_n < K_1$, use treatment A on the remaining $N-2n$;

If $K_1 \leqq u_n \leqq K_2$, use each treatment on the half of the remaining,

where $2n$ is the number of individuals for the trial and where K_1 and K_2 are assigned real values and $K_1 < K_2$.

The Second Case: Perform a trial for n drawn from N individuals on the treatment A, where the parameter θ_A associated with the treatment A is unknown to us and the value of the parameter θ_B associated with the treatment B known.

Compute the observed value of $u_n = \sum_{i=1}^{n} v_i$.

Procedure: If $u_n > K_2$, use treatment B on the remaining $N-n$;

If $u_n < K_1$, use treatment A on the remaining $N-n$;

If $K_1 \leqq u_n \leqq K_2$, use each treatment on the half of the remaining,

where n is the number of individuals for the trial and where K_1 and K_2 are assigned real values and $K_1 < K_2$.

The optimum values n^* (or $p^* = n^*/N$), K_1^*, and K_2^* of n (or $p = n/N$), K_1, and K_2 should be determined respectively so that the over-all expected loss (or the over-all expected net gain) constructed in section **1.2** is minimized (or maximized) under our procedure.

§1.2. Construction of Over-all Expected Loss Function.

The First Case: When the true value of the parameter is δ, if $\delta \in \Omega_0$, then the expected loss $[E\ Loss]_A$ for N individuals is obtained as follows by **Definition 2.**

(1.1) $$[E\ Loss]_A = C(\delta^* - \delta) \Big[n + \{N-2n\} \cdot P_r\{u_n < K_1\} + \frac{N-2n}{2} \cdot P_r\{K_1 \leqq u_n \leqq K_2\} \Big].$$

For the same reason, if $\delta \in \Omega_1$, then the expected loss $[E\ Loss]_B$ for N individuals

is obtained as follows.

(1.2) $[E\ Loss]_B = C(\delta - \delta^*)\left[n + \{N - 2n\} \cdot P_r\{u_n > K_2\} + \dfrac{N - 2n}{2} \cdot P_r\{K_1 \leqq u_n \leqq K_2\} \right].$

By integrating (1.1) and (1.2) over the *a priori* distribution for the parameter δ, the over-all expected loss denoted by $\overline{E\ Loss}$ is then obtained as follows.

(1.3) $\overline{E\ Loss}/NC = \displaystyle\int_{\delta^*}^{\Delta_2} (\delta - \delta^*)g(\delta)d\delta + p\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta)g(\delta)d\delta$

$+ (1 - 2p)\displaystyle\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta) \cdot P_r\{u_n < K_1\}g(\delta)d\delta$

$+ \left(\dfrac{1}{2} - p\right)\displaystyle\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta) \cdot P_r\{K_1 \leqq u_n \leqq K_2\}g(\delta)d\delta\ ,$

where $p = n/N$ and $g(\delta)$ denotes the probability density function of the *a priori* distribution for δ.

The Second Case: In the same manner, we have $[E\ Loss]_A$, $[E\ Loss]_B$, and $\overline{E\ Loss}$.

(1.1′) $[E\ Loss]_A = C(\delta^* - \delta)\left[n + (N - n) \cdot P_r\{u_n < K_1\} + \dfrac{N - n}{2} \cdot P_r\{K_1 \leqq u_n \leqq K_2\} \right],$

(1.2′) $[E\ Loss]_B = C(\delta - \delta^*)\left[(N - n) \cdot P_r\{u_n > K_2\} + \dfrac{N - n}{2} \cdot P_r\{K_1 \leqq u_n \leqq K_2\} \right],$

(1.3′) $\overline{E\ Loss}/NC = \displaystyle\int_{\delta^*}^{\Delta_2}(\delta - \delta^*)g(\delta)d\delta$

$+ p\displaystyle\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta)g(\delta)d\delta + (1 - p)\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta) \cdot P_r\{u_n < K_1\}g(\delta)d\delta$

$+ \dfrac{1 - p}{2}\displaystyle\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta) \cdot P_r\{K_1 \leqq u_n \leqq K_2\}g(\delta)d\delta\ ,$

where $p = n/N$.

We should like to determine the optimum values p^*, K_1^* and K_2^* of p, K_1 and K_2 respectively so that (1.3) is minimized in the first case and (1.3′) is minimized in the second case.

§1.3. Construction of Over-all Expected Net Gain Function.

The First Case: Let δ be the true value of the parameter. If $\delta \in \Omega_0$, then the expected net gain is obtained as follows by **Definition 3**.

(1.4) $E\ Net\ Gain = G(\delta^* - \delta)(N - 2n)\{P_r(\text{Select }B) - P_r(\text{Select }A)\}\ .$

If $\delta \in \Omega_1$, then the expected net gain is obtained.

(1.5) $E\ Net\ Gain = G(\delta - \delta^*)(N - 2n)\{P_r(\text{Select }A) - P_r(\text{Select }B)\}\ .$

Therefore, the over-all expected net gain is obtained.

(1.6) $\overline{E\ Net\ Gain}/NG = (1 - 2p)\left[\displaystyle\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta)g(\delta)d\delta \right.$

$-2\displaystyle\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta) \cdot P_r\{u_n < K_1\}g(\delta)d\delta$

$\left. -\displaystyle\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta) \cdot P_r\{K_1 \leqq u_n \leqq K_2\}g(\delta)d\delta \right],$

where $p = n/N$, and where $g(\delta)$ is the probability density function in the *a priori* distribution for δ.

From (1.3) and (1.6), we then get the relation between $\overline{E\ Loss}/NC$ and $\overline{E\ Net\ Gain}/NG$.

(1.7) $$\overline{E\ Net\ Gain}/NG = [\text{Constant}] - 2[\overline{E\ Loss}/NC].$$

The Second Case: Let δ be the true value of the parameter. If $\delta \in \Omega_0$, then the expected net gain becomes

(1.4′) $$E\ Net\ Gain = G(\delta^* - \delta)[-n + (N-n) \cdot \{P_r(\text{Select } B) - P_r(\text{Select } A)\}].$$

If $\delta \in \Omega_1$, then the expected net gain becomes

(1.5′) $$E\ Net\ Gain = G(\delta - \delta^*)[n + (N-n) \cdot \{P_r(\text{Select } A) - P_r(\text{Select } B)\}].$$

Therefore, the over-all expected net gain is obtained.

(1.6′) $$\overline{E\ Net\ Gain}/NG = (1-2p)\int_{\Delta_1}^{\Delta_2} (\delta^* - \delta)g(\delta)d\delta$$

$$-2(1-p)\int_{\Delta_1}^{\Delta_2} (\delta^* - \delta) \cdot P_r\{u_n < K_1\}g(\delta)d\delta$$

$$-(1-p)\int_{\Delta_1}^{\Delta_2} (\delta^* - \delta) \cdot P_r\{K_1 \leq u_n \leq K_2\}g(\delta)d\delta,$$

where $g(\delta)$ is the probability density function in the *a priori* distribution for δ.

From (1.3′) and (1.6′), we then get the relation between $\overline{E\ Loss}/NC$ and $\overline{E\ Net\ Gain}/NG$.

(1.7′) $$\overline{E\ Net\ Gain}/NG = [\text{Constant}] - 2[\overline{E\ Loss}/NC].$$

We should like to determine the optimum values p^*, K_1^* and K_2^* of p, K_1 and K_2 respectively so that (1.6) is maximized in the first case and (1.6′) is maximized in the second case.

§1.4. Examples.

Example (1): Normal Distribution.

(1°) **The First Case:** In the first place, let us consider the case when both population means μ_A and μ_B of two normal distributions are unknown to us and both population variances are common (denoted by σ^2) and known to us.

When y_i's $(i = 1, 2, 3, \cdots, n)$ are drawn independently from the normal distribution with an unknown mean $\delta = (\mu_A - \mu_B)/(\sqrt{2}\sigma)$ and unite variance, $u_n = -\sum_{i=1}^{n} y_i$ is the sufficient statistic of the parameter δ for this class of distributions with the probability density function

(1.8) $$f_n(u_n; \delta)d\mu(u_n) = exp\left\{-\delta u_n - n \cdot \frac{\delta^2}{2} - \frac{u_n^2}{2n} + log\,(2\pi n)^{-\frac{1}{2}}\right\}d\mu(u_n)$$

$(n = 1, 2, 3, \cdots)$, where $b(\delta) = -\delta^2/2$ and $\delta^* = 0$.

We assume that an *a priori* distribution for the parameter δ has a normal distribution with zero mean and known variance $\sigma_0^2/(2\sigma^2)$, where $\mathbf{I} = (-\infty, \infty)$. Let us suppose that $K_1 = -K$ and $K_2 = K$, where K is a positive real number.

We then have

(1.9) $$P_r\{u_n < -K\} = 1 - \Phi\left(\frac{K - n\sigma' z}{\sqrt{n}}\right),$$

(1.10) $$P_r\{-K \leqq u_n \leqq K\} = \Phi\left(\frac{K - n\sigma' z}{\sqrt{n}}\right) - \Phi\left(\frac{-K - n\sigma' z}{\sqrt{n}}\right),$$

where $\Phi(t) = \int_{-\infty}^{t} \varphi(u)du$ and $\varphi(u) = (2\pi)^{-\frac{1}{2}} exp\{-u^2/2\}$, and where $\sigma' = \sigma_0/(\sqrt{2}\sigma)$.

Therefore, substituting (1.9) and (1.10) into (1.6), we have

(1.11) $$\overline{E\ Net\ Gain}/NG = -\sigma'(1 - 2p)\left[\int_{-\infty}^{\infty} z\Phi\left(\frac{K - n\sigma' z}{\sqrt{n}}\right)\varphi(z)dz\right.$$
$$\left. + \int_{-\infty}^{\infty} z\Phi\left(\frac{-K - n\sigma' z}{\sqrt{n}}\right)\varphi(z)dz\right].$$

The optimum values K^* and p^* of K and p which maximize $\overline{E\ Net\ Gain}/NG$ of (1.11) have been shown in the previous papers [1] and [4], that is, $K^* = 0$ and $p^* = 1/(3 + \sqrt{9 + 4R})$, where $R = \sigma'^2 N = (\sigma_0^2 N)/(2\sigma^2)$. Moreover, the maximum values of $\overline{E\ Net\ Gain}/NG$ have been computed at various values of R.

(2°) **The Second Case:** In the next place, let us consider the case when one population mean μ_A is unknown to us and another population mean μ_B is known.

When y_i's $(i = 1, 2, 3, \cdots,)$ are drawn independently from the normal distribution with an unknown mean $\delta = (\mu_A - \mu_B)/\sigma$ and unite variance, $u_n = -\sum_{i=1}^{n} y_i$ is the sufficient statistic of the parameter δ for this class of distributions with the probability density function such that $b(\delta) = -\delta^2/2$, $a_n(u_n) = -u_n^2/(2n) + log(2\pi n)^{-\frac{1}{2}}$, and $\delta^* = 0$.

We assume that the *a priori* distribution for the parameter δ has a normal distribution with zero mean and known variance σ_0^2/σ^2, where $\mathbf{I} = (-\infty, \infty)$. Let us suppose that $K_1 = -K$ and $K_2 = K$, where K is a positive real number.

In the previous paper [9], we have discussed in detail on this case. The optimum values K^* and p^* of K and p which maximize (1.6′) have been given there, that is, $K^* = 0$ and $p^* = 2/(3 + \sqrt{9 + 16R})$, where $R = (N\sigma_0^2)/(2\sigma^2)$. Moreover, the maximum values of the over-all expected net gain have been computed at various values of R.

Example (2): Binomial Distribution.

(1°) **The First Case:** Let us firstly consider the case when both population proportions (i. e., effective proportions) $\mathfrak{p}_A$ and $\mathfrak{p}_B$ of two binomial distributions are unknown to us.

Let us drawn $2n$ from N individuals in the form of pair and perform treatment A on one member of the pair and treatment B on the other.

Let $a_1, \cdots, a_n$ be the results based on performing treatment A, and $b_1, \cdots, b_n$ the results based on performing treatment B, where a_i and b_i are value 1 or 0, and where value 1 denotes that the performed treatment j is effective and value 0 denotes that the performed treatment j is ineffective $(i = 1, 2, \cdots, n; j = A, B)$. These results are arranged in the order observed. Thus we can consider the sequence of n pairs:

$$(a_1, b_1), \cdots, (a_n, b_n).$$

Let t_1 be the number of pairs $(1, 0)$ and t_2 the number of pairs $(0, 1)$ in this sequence and let $t = t_1 + t_2$. We consider only the pairs $(0, 1)$ and $(1, 0)$.

Since each pair (a_i, b_i) is equal to one of the pairs $(0, 1)$ and $(1, 0)$, the conditional probability that it is equal to $(0, 1)$ is given by

$$\mathfrak{p} = \frac{(1-\mathfrak{p}_A)\mathfrak{p}_B}{\mathfrak{p}_A(1-\mathfrak{p}_B)+(1-\mathfrak{p}_A)\mathfrak{p}_B}$$

and the conditional probability that it is to equal to $(1, 0)$ is given by $1-\mathfrak{p}$.

When y_i's $(i = 1, 2, 3, \cdots, t)$ are drawn independently from the binomial population with an unknown population proportion $\mathfrak{p}$, $u_t = \sum_{i=1}^{t} y_i$ is the sufficient statistic of the parameter $\delta = \log \mathfrak{p}^{-1}(1-\mathfrak{p}) = -\log u$ for this class of distributions with the probability density function

$$f_t(u_t ; \delta) = exp \{-\delta u_t + tb(\delta) + a_t(u_t)\}$$

$(t = 1, 2, \cdots)$, where $b(\delta) = -\log(1+e^{-\delta})$ and $a_t(u_t) = -\log[(t+1) \cdot B(t-u_t+1, u_t+1)]$ and where u denotes an odds ratio $(\mathfrak{p}_B(1-\mathfrak{p}_A))/(\mathfrak{p}_A(1-\mathfrak{p}_B))$.

Let us define $\delta^* = 0$. We have then that the acceptance of $\delta \in \Omega_0$ is equivalent to the acceptance of the event "$1 < u$" and the acceptance of $\delta \in \Omega_1$ is equivalent to the acceptance of the event "$1 > u$".

Let us assume that the *a priori* distribution for the parameter δ has a uniform distribution $1/(\Delta_2 - \Delta_1)$, where $\Delta_i = -\log d_i$ $(i = 1, 2)$ $(-\infty < \Delta_1 < 0 < \Delta_2 < \infty)$, and the interval $\langle d_2, d_1 \rangle$ denotes an *a priori* defined interval for odds ratio u $(0 < d_2 < 1 < d_1 < \infty)$.

We can now compute the over-all expected loss $\overline{E\ Loss}/NC$ of (1.3) as follows.

(1.12)
$$\overline{E\ Loss}/NC = \frac{1}{\Delta_2 - \Delta_1} \left\{ \frac{1}{2} \{(1-p)\Delta_2^2 + p\Delta_1^2\} \right.$$

$$-(1-2p) \int_{\Delta_1}^{\Delta_2} \delta \cdot P_r\{u_t < K_1\} d\delta$$

$$\left. -(0.5-p) \int_{\Delta_1}^{\Delta_2} \delta \cdot P_r\{K_1 \leq u_t \leq K_2\} d\delta \right\}.$$

We have discussed on this case in the previous paper [11] and computed the optimum values p^* and the minimum values of $\overline{E\ Loss}/NC$ with respect to various types of *a priori* distribution for δ, where we define $K_1 = -K$ and $K_2 = K$ and K is a positive real value.

(2°) **The Second Case:** In the next place, let us consider the case when one population proportion (i. e., effective proportion) $\mathfrak{p}_A$ is unknown to us and the other $\mathfrak{p}_B$ known.

When y_i's $(i = 1, 2, \cdots,)$ are drawn independently from the binomial population with an unknown population proportion $\mathfrak{p}_A$, $u_n = \sum_{i=1}^{n} y_i$ is the sufficient statistic of the parameter $\delta = \log \mathfrak{p}_A^{-1}(1-\mathfrak{p}_A)$ for this class of distributions with the probability density function such that $b(\delta) = -\log(1+e^{-\delta})$, $a_n(u_n) = -\log[(n+1) \cdot B(n-u_n+1, u_n+1)]$, and $\delta^* = \log \mathfrak{p}_B^{-1}(1-\mathfrak{p}_B)$.

In the previous paper [6], we have been in detail concerned with the case such

that $K_1 = K_2 = n_0$, where n_0 indicates the maximum nonnegative integer $[n \cdot \mathfrak{p}_B]$ such that is not larger than $n \cdot \mathfrak{p}_B$.

Example (3): Gamma Distribution $\Gamma(\alpha, \delta)$.

Let us consider the case when v_i's $(i = 1, 2, \cdots, m)$ are drawn independently from the Γ-distribution $\Gamma(\alpha, \delta)$ with unknown mean α/δ, where $0 < v_i < \infty$, $0 < \delta < \infty$, and $0 < \alpha < \infty$. Then $u_m = \sum_{i=1}^{m} v_i$ is the sufficient statistic for the parameter δ with the probability density function

$$(1.13) \qquad f_m(u_m; \delta) d\mu(u_m) = exp\{-\delta u_m + m\alpha \, log \, \delta + log\,(\Gamma(m\alpha)^{-1} \cdot u_m^{m\alpha-1})\} d\mu(u_m),$$

where

$$(1.14) \qquad b(\delta) = \alpha \, log \, \delta, \qquad a_m(u_m) = log\,(\Gamma(m\alpha)^{-1} \cdot u_m^{m\alpha-1}),$$

and where m is any positive integer.

When an *a priori* distribution for the parameter δ is given, the over-all expected loss can be computed from (1.3) or (1.3').

Example (4): Poisson Distribution $P(\lambda)$.

Let us consider the case when v_i's $(i = 1, 2, \cdots, m)$ are drawn independently from the Poisson distribution $P(\lambda)$, where $v_i = 0, 1, 2 \cdots$, (nonnegative integer) and $0 < \lambda < \infty$. Then $u_m = \sum_{i=1}^{m} v_i$ is the sufficient statistic for this class of distributions and its probability density function is

$$(1.15) \qquad f_m(u_m; \delta) d\mu(u_m) = exp\{-\delta u_m - m e^{-\delta} + log\,(m^{u_m} \cdot \Gamma(u_m+1)^{-1})\} d\mu(u_m)$$

defined over $0 \leq u_m < \infty$, where

$$(1.16) \qquad \delta = log\,(1/\lambda), \qquad b(\delta) = -e^{-\delta}, \quad \text{and} \quad a_m(u_m) = log\,(m^{u_m} \cdot \Gamma(u_m+1)^{-1}),$$

and where μ is a step-function having jumps of height 1 at every nonnegative integer u_m.

When an *a priori* distribution for the parameter δ is given, the over-all expected loss can be computed from (1.3) or (1.3').

CHAPTER 2

SEQUENTIAL PLAN IN A POPULATION WITH ONE PARAMETER EXPONENTIAL DISTRIBUTION

§2.1. Some Preparations from a Sequential Probability Ratio Test by A. Wald [13].

It is necessary to compute OC function $L(\delta)$ and ASN function $E_\delta(n)$ for the purpose of constructing the expected loss function in section **2.3**. Therefore, we shall give some preparations concerning a sequential probability ratio test for a population with one parameter exponential distribution.

We hereupon assume particularly with respect to the common measure μ in (0.1) and (0.2) that there exists at least one pair (δ_0, δ_1) of δ such that

$$(2.1) \qquad \mu\left(\left(\frac{b(\delta_1) - b(\delta_0)}{\delta_1 - \delta_0}, +\infty\right)\right) > 0,$$

where $\varDelta_1 \leqq \delta_0 < \delta_1 \leqq \varDelta_2$.

Let us consider only the pair (δ_0, δ_1) of δ throughout this section such that satisfies assumption (2.1).

Let us consider the probability ratio

(2.2) $$p_{1m}/p_{0m} = \prod_{i=1}^{m} \frac{f(v_i; \delta_1)}{f(v_i; \delta_0)} = exp\{-(\delta_1 - \delta_0)u_m + m\{b(\delta_1) - b(\delta_0)\}\},$$

where we consider the sequential probability ratio test of strength (α, β) for testing the hypothesis H_0 that $\delta = \delta_0$ against the alternative H_1 that $\delta = \delta_1$ $(\delta_0 < \delta_1)$.

Whence, letting

(2.3) $$z_i = log\frac{f(v_i; \delta_1)}{f(v_i; \delta_0)} = -(\delta_1 - \delta_0)v_i + b(\delta_1) - b(\delta_0) \quad \text{and} \quad Z_m = \sum_{i=1}^{m} z_i,$$

we obtain

(2.4) $$Z_m = log(p_{1m}/p_{0m}) = -(\delta_1 - \delta_0)u_m + m\{b(\delta_1) - b(\delta_0)\}.$$

According to the general approach in the sequential probability ratio test, there are two constants A and B to decide one of the following three alternatives:

(1°) When we have

(2.5) $$\begin{cases} Z_m \geqq log\,A, \\ \text{that is,} \\ u_m \leqq \dfrac{1}{\delta_0 - \delta_1}\,log\,A + m \cdot \dfrac{b(\delta_1) - b(\delta_0)}{\delta_1 - \delta_0}, \end{cases}$$

then we stop the experimentation and accept the hypothesis H_1.

(2°) When we have

(2.6) $$\begin{cases} log\,B < Z_m < log\,A, \\ \text{that is,} \\ \dfrac{1}{\delta_0 - \delta_1}\,log\,B + m \cdot \dfrac{b(\delta_1) - b(\delta_0)}{\delta_1 - \delta_0} > u_m > \dfrac{1}{\delta_0 - \delta_1}\,log\,A + m \cdot \dfrac{b(\delta_1) - b(\delta_0)}{\delta_1 - \delta_0}, \end{cases}$$

then we make another independent statistic v_{m+1} which gives us $u_{m+1} = \sum_{i=1}^{m+1} v_i$.

(3°) When we have

(2.7) $$\begin{cases} Z_m \leqq log\,B, \\ \text{that is,} \\ u_m \geqq \dfrac{1}{\delta_0 - \delta_1}\,log\,B + m \cdot \dfrac{b(\delta_1) - b(\delta_0)}{\delta_1 - \delta_0}, \end{cases}$$

then we stop the experimentation and accept H_0.

Therefore, the two parallel straight lines L_1 and L_0 giving the boundaries for a continuation of experimentations are given by

(2.8) $$\begin{cases} L_1: \quad r_m = \dfrac{b(\delta_1) - b(\delta_0)}{\delta_1 - \delta_0} \cdot m + \dfrac{1}{\delta_0 - \delta_1}\,log\,B, \\ L_0: \quad a_m = \dfrac{b(\delta_1) - b(\delta_0)}{\delta_1 - \delta_0} \cdot m + \dfrac{1}{\delta_0 - \delta_1}\,log\,A. \end{cases}$$

Thus the common slope of the lines L_1 and L_0 is given by

(2.8')
$$s = \frac{b(\delta_1) - b(\delta_0)}{\delta_1 - \delta_0} \ .$$

The intercept of L_1 is equal to

(2.8'')
$$h_1 = \frac{1}{\delta_0 - \delta_1} \log \boldsymbol{B}$$

and the intercept of L_0 is given by

(2.8''')
$$h_0 = \frac{1}{\delta_0 - \delta_1} \log \boldsymbol{A}, \quad \text{where} \quad A \sim \frac{1-\beta}{\alpha} \quad \text{and} \quad B \sim \frac{\beta}{1-\alpha} \ .$$

The essential instrument for constructing the fundamental aspects as well as the practical calculations concerning sequential analysis is to find out a function $h(\delta)$ such that for each value δ

$$E_{\delta}\Big[\Big(\frac{f(v\,;\,\delta_1)}{f(v\,;\,\delta_0)}\Big)^{h(\delta)}\Big] = 1\,,$$

that is,

(2.9)
$$E_{\delta}[exp\,\{zh(\delta)\}] = 1\,,$$

where

$$z = \log \frac{f(v\,;\,\delta_1)}{f(v\,;\,\delta_0)} = -(\delta_1 - \delta_0)v + b(\delta_1) - b(\delta_0)\,,$$

and that $h(\delta) \neq 0$ for each δ.

Theorem 1. (Wald [13], p. 158) *Let z be a random variable such that the following three conditions are fulfilled:*

(I) *The expected value $E(z)$ exists and is not equal to 0.*

(II) *There exists a positive ε such that*

$$P_r\{e^z < 1 - \varepsilon\} > 0 \quad and \quad P_r\{e^z > 1 + \varepsilon\} > 0\,.$$

(III) *For any real value h the expected value $E(e^{hz}) = g(h)$ exists.*
Then there exists one and only one real value $h^* \neq 0$ such that

$$E(e^{h^*z}) = 1\,.$$

Let us denote by δ' a specific value of δ satisfying

(2.10)
$$b'(\delta) = \frac{b(\delta_1) - b(\delta_0)}{\delta_1 - \delta_0}$$

for assigned values δ_0 and δ_1 of δ.

It should be noted that there exists one and only one δ' satisfying (2.10) and moreover $\delta_0 < \delta' < \delta_1$ since $b(\delta)$ has the first derivative $b'(\delta)$ and $b'(\delta)$ is continuous and strictly monotone decreasing in $\mathbf{I}$ from our conditions.

Let us denote by $\mathbf{I}'$ the difference set $\mathbf{I} - \{\delta'\}$.

Theorem 2. *Let $\{z_i\}$ be a set of independent random variables defined by (2.3). Then each random variable z_i in a set of $\{z_i\}$ satisfies the three conditions (I), (II), and (III) enunciated in Theorem 1 for each δ belonging to $\mathbf{I}'$.*

Proof. Ad (I). Due to our condition with respect to $b(\delta)$, we have for each z_i

$$E(z_i) = -(\delta_1 - \delta_0)b'(\delta) + b(\delta_1) - b(\delta_0) \,.$$

Moreover, we have $E(z_i) \neq 0$ for any δ belonging to $\mathbf{I}'$.

Ad (II). Let ε be any assigned value in $0 < \varepsilon < 1$.

Let us assume ad absurdity that

$$P_r\{exp\,\{z_i\} < 1-\varepsilon\} = P_r\Big\{v_i > \frac{b(\delta_1) - b(\delta_0) - log\,(1-\varepsilon)}{\delta_1 - \delta_0}\Big\}$$

be equal to 0.

But, since we consider only the pair (δ_0, δ_1) satisfying (2.1), as we mentioned earlier, we have

$$\int_{\zeta(\delta_0, \delta_1)}^{\zeta(\delta_0, \delta_1;\varepsilon)} f(v_i ; \delta)d\mu(v_i) + \int_{\zeta(\delta_0, \delta_1;\varepsilon)}^{\infty} f(v_i ; \delta)d\mu(v_i) = \int_{\zeta(\delta_0, \delta_1)}^{\infty} f(v_i ; \delta)d\mu(v_i) > 0 \,,$$

where

$$\zeta(\delta_0, \delta_1) = \{b(\delta_1) - b(\delta_0)\}/(\delta_1 - \delta_0) \,,$$

$$\zeta(\delta_0, \delta_1 ; \varepsilon) = \{b(\delta_1) - b(\delta_0)\}/(\delta_1 - \delta_0) - \{log\,(1-\varepsilon)\}/(\delta_1 - \delta_0) \,,$$

$$\text{and} \quad \lim_{\varepsilon \to +0} \zeta(\delta_0, \delta_1 ; \varepsilon) = \zeta(\delta_0, \delta_1) \,.$$

Therefore,

$$\lim_{\varepsilon \to +0} \int_{\zeta(\delta_0, \delta_1;\varepsilon)}^{\infty} f(v_i ; \delta)d\mu(v_i) = \int_{\zeta(\delta_0, \delta_1)}^{\infty} f(v_i ; \delta)d\mu(v_i) > 0 \,,$$

and whence the assumption of our absurdity contradicts to this.

Similarly, we can verify the second part of (II).

Ad (III). For any real value h, we have

$$g(h) = E(e^{hz}) = exp\,\{\{b(\delta_1) - b(\delta_0)\}h + b(\delta)\} \cdot exp\,\{-b(\delta^{(*)})\} \,,$$

where $\delta^{(*)} = (\delta_1 - \delta_0)h + \delta$.

Therefore, $g(h)$ exists for any real value h.

The immediate consequence of Theorem 1 and Theorem 2 is the following.

There exists exactly a function $h(\delta)$ satisfying (2.9) for each value of δ belonging to $\mathbf{I}'$.

Therefore, on the basis of the theorem **7.1** by Kitagawa [3], the function $h(\delta)$ defined over $\mathbf{I}'$ can be computed by solving the following equation

$$(2.11) \qquad b(h(\delta) \cdot (\delta_1 - \delta_0) + \delta) = h(\delta) \cdot (b(\delta_1) - b(\delta_0)) + b(\delta) \,,$$

where it should be noted that the form of $b(\delta)$ is actually given to us.

When the excess of p_{1m}/p_{0m} over the boundaries $\mathbf{A}$ and $\mathbf{B}$ at the termination of the sequential process is neglected, we get the approximation formula:

$$(2.12) \qquad L(\delta) \sim \frac{exp\,\{h(\delta)\,log\,\mathbf{A}\} - 1}{exp\,\{h(\delta)\,log\,\mathbf{A}\} - exp\,\{h(\delta)\,log\,\mathbf{B}\}} \,,$$

where $L(\delta)$ is defined as the probability that the sequential process will terminate with the acceptance of H_0 when δ is the true value of the parameter.

From (2.8$'$), we also obtain

$$(2.12') \qquad L(\delta) \sim \frac{exp\,\{h(\delta) \cdot (\delta_0 - \delta_1)h_0\} - 1}{exp\,\{h(\delta) \cdot (\delta_0 - \delta_1)h_0\} - exp\,\{h(\delta) \cdot (\delta_0 - \delta_0)h_1\}} \,.$$

Since $E_\delta(z) = -(\delta_1 - \delta_0) \cdot b'(\delta) + b(\delta_1) - b(\delta_0) \neq 0$ for any δ belonging to $\mathbf{I}'$, we obtain similarly

(2.13) $$E_\delta(n) \sim \frac{L(\delta)\, log\, \mathbf{B} + \{1 - L(\delta)\}\, log\, \mathbf{A}}{E_\delta(z)} \;,$$

where n denotes the number of observed pairs required by the test and $E_\delta(z)$ denotes the expected value of z when δ is the true value of the parameter.

Therefore, in like (2.11'),

(2.13') $$E_\delta(n) \sim \frac{(h_1 - h_0) \cdot L(\delta) + h_0}{b'(\delta) - s} \;,$$

where $L(\delta)$ is given by (2.12').

When δ is equal to δ' satisfying a relation (2.10), let us make use of the following limiting values $\lim_{\delta \to \delta'} L(\delta)$ and $\lim_{\delta \to \delta'} E_\delta(n)$ instead of $L(\delta')$ and $E_{\delta'}(n)$ respectively.

(2.12'') $$L(\delta') \equiv \lim_{\delta \to \delta'} L(\delta) \sim \frac{log\, \mathbf{A}}{log\, \mathbf{A} - log\, \mathbf{B}} = \frac{h_0}{h_0 - h_1} \;,$$

(2.13'') $$E_{\delta'}(n) \equiv \lim_{\delta \to \delta'} E_\delta(n) \sim \frac{-log\, \mathbf{A} \cdot log\, \mathbf{B}}{E_\delta(z^2)} = \frac{h_0 h_1}{b''(\delta')} \;.$$

§ 2.2. Procedure.

We now suppose that the trial no longer calls for a fixed number of participants, but the trial is performed sequentially on a pair of individuals at a time, treatment A on one member of the pair and treatment B on the other. After the results from each pair are available, a decision based on the cumulative evidence at hand is made to select one of the two treatments as the superior and use it on the all remaining individuals or to continue the trial by having an additional pair participate.

Procedure. After the m-th pair, compute the value of a cumulative sum u_m of statistics $v_1, v_2, \cdots, v_m$.

If $u_m \geqq K + s \cdot m$, then select treatment B and perform B on the
remaining individuals;

(2.14)   If $u_m \leqq -K + s \cdot m$, then select treatment A and perform A on the
remaining individuals;

If $-K + s \cdot m < u_m < K + s \cdot m$, then continue with another pair,

where K is a positive real number and s is a real number, and where K and $-K$ are the intercepts of the two parallel straight lines indicating the location of boundaries at the termination of our sequential process and s is a common slope of these two straight lines.

The problem is to determine the optimum values K^* and s^* of K and s respectively so that the expected loss constructed in the following section **2.4** is minimized.

§ 2.3. Construction of Over-all Expected Loss Function.

Let us denote by $P_r\{$Select treatment $B | \delta\}$ the probability that we select treatment B as the superior at the terminal decision of our sequential process when δ is

the true value of the parameter. We have then

(2.15) $\qquad P_r\{\text{Select treatment } B \,|\, \delta\} = P_r\{u_m \geqq K + s \cdot m \,|\, \delta\}$

$\qquad\qquad\qquad = P_r\{u_m \geqq h_1 + s \cdot m \,|\, \delta\} = P_r\{Z_m \leqq \log \boldsymbol{B} \,|\, \delta\} = L(\delta)\,.$

Therefore, we have similarly

(2.15′) $\qquad\qquad\qquad P_r\{\text{Select treatment } A \,|\, \delta\} = 1 - L(\delta)\,,$

where $L(\delta)$ is given by (2.11′) or (2.11″).

Assuming N to be large so that the formula for unrestricted (i. e., open) sequential sampling are reasonable approximations, we obtain the following expected loss function.

If treatment A is truly inferior to B (i. e., $\delta \in \varOmega_0$), then the expected loss $[E\ Loss]_A$ for N individuals is obtained as follows, by **Definition 2** and the abovementioned procedure.

(2.16) $\qquad [E\ Loss]_A = C(\delta^* - \delta)[E_{\delta}(n) + \{N - 2E_{\delta}(n)\} \cdot \{1 - L(\delta)\}]\,.$

On the other hand, if treatment B is truly inferior to A (i. e., $\delta \in \varOmega_1$), then the expected loss $[E\ Loss]_B$ for N individuals is obtained as follows.

(2.17) $\qquad [E\ Loss]_B = C(\delta - \delta^*)[E_{\delta}(n) + \{N - 2E_{\delta}(n)\} \cdot L(\delta)]\,.$

Consequently, by integrating over the *a priori* distribution for δ based on the asumption, the over-all expected loss $\overline{E\ Loss}$ is obtained as follows.

(2.18) $\qquad \overline{E\ Loss}/NC = \int_{\varDelta_1}^{\delta^*} (\delta^* - \delta)g(\delta)d\delta - \frac{1}{N}\int_{\varDelta_1}^{\varDelta_2}(\delta^* - \delta)E_{\delta}(n)g(\delta)d\delta$

$\qquad\qquad\qquad - \int_{\varDelta_1}^{\varDelta_2}(\delta^* - \delta)L(\delta)g(\delta)d\delta + \frac{2}{N}\int_{\varDelta_1}^{\varDelta_2}(\delta^* - \delta)L(\delta)E_{\delta}(n)g(\delta)d\delta\,,$

where $g(\delta)$ denotes the probability density function of the *a priori* distribution for δ.

$L(\delta)$ in (2.12′) and $E_{\delta}(n)$ in (2.13′) are transformed by our procedure respectively as follows.

(2.19) $\qquad\qquad\qquad L(\delta) \sim \dfrac{1}{1 + e^{(\delta_0 - \delta_1)h(\delta)K}}$

and

(2.20) $\qquad\qquad\qquad E_{\delta}(n) \sim \dfrac{K\{2L(\delta) - 1\}}{b'(\delta) - s}\,,$

where $h(\delta)$ can be obtained by (2.10).

Giving actually the forms of $b(\delta)$, $h(\delta)$ and $g(\delta)$, we can now compute $\overline{E\ Loss}/NC$ of (2.18).

In conclusion, we should ask for the optimum values K^* and s^* of K and s respectively so that is (2.18) is minimized.

The logical relation between the sequential probability ratio test by Wald [13] and our sequential scheme can be enunciated as follows.

First of all, we have assigned a positive intercept K and a common slope s of the two parallel straight lines indicating the location of boundaries at the termination of our sequential process.

Let us suppose that values of δ_0 and δ_1 (, where δ_0 and δ_1 satisfy (2.1),) were

assigned. We then could determine approximately values of α and β on the basis of (2.8′). For these assigned quantities α, β, δ_0, and δ_1, we could perform the sequential probability ratio test of strength (α, β) for testing the hypothesis H_0 that $\delta = \delta_0$ against the alternative H_1 that $\delta = \delta_1$. Therefore, we could compute OC function $L(\delta)$ and consequently construct $\overline{E\ Loss}$ due to our assumption.

It has been our problem to determine specific values K^* and s^* of K and s respectively so that $\overline{E\ Loss}$ is minimized.

Thus we should actually compute specific values α^* and β^* of α and β corresponding to K^* and s^* once more on the basis of (2.8′), for firstly assigned value of δ_0 and δ_1, and then obtain a specific OC function $L(\delta)$ corresponding to these values of α^*, β^*, δ_0, and δ_1.

We have been so far concerned with the third case in which the two parameters θ_A and θ_B respectively associated with the two treatments A and B are both unknown to us. However, the fourth case in which the parameter θ_B is known to us but the other θ_A is unknown to us can be discussed in the same manner. By performing successively the only one treatment (i. e., treatment A) with unknown parameter on each individual randomly drawn from N individuals, we con then carry out our sequential process.

In this case, the considered statistic u_m indicates a discrepancy between the cumulative observations from performing the treatment with unknown parameter and the known value of the parameter of the other treatment.

We can now construct the over-all expected loss as follows on the basis of **Definition 2** on the loss.

Let δ be the true value of the parameter. First of all, if $\delta \in \Omega_0$, then the expected loss $[E\ Loss]_A$ for N individuals becomes as follows.

(2.21) $[E\ Loss]_A = C(\delta^* - \delta)[E_{\delta}(n) + \{N - E_{\delta}(n)\} \cdot \{1 - L(\delta)\}]$.

In the second place, if $\delta \in \Omega_1$, then the expected loss $[E\ Loss]_B$ over the whole trials becomes as follows.

(2.22) $[E\ Loss]_B = C(\delta - \delta^*)\{N - E_{\delta}(n)\} \cdot L(\delta)$.

The over-all expected loss denoted by $\overline{E\ Loss}$ is finally as follows.

(2.23) $$\overline{E\ Loss}/NC = \int_{\Delta_1}^{\delta^*}(\delta^* - \delta)g(\delta)d\delta - \int_{\Delta_1}^{\Delta_2}(\delta^* - \delta)L(\delta)g(\delta)d\delta$$
$$+ \frac{1}{N}\int_{\Delta_1}^{\Delta_2}(\delta^* - \delta)L(\delta)E_{\delta}(n)g(\delta)d\delta ,$$

where $L(\delta)$ and $E_{\delta}(n)$ can be obtained from (2.19) and (2.20) respectively, and where $g(\delta)$ denotes the probability density function of the a priori distribution for δ.

In conclusion, we should ask for the optimum values K^* and s^* of K and s respectively so that (2.23) is minimized.

§ 2.4. Examples.

Example 1: Normal Distribution with Known Variance.
(1°) **The Third Case:** We are firstly concerned with the case when population

means μ_A and μ_B of two normal populations with known and common variance σ^2 are both unknown to us.

When y_i's $(i = 1, 2, \cdots, m)$ are drawn independently from the normal population with unknown population mean $\delta = (\mu_A - \mu_B)/(\sqrt{2}\sigma)$ and unite variance, $u_m = \sum_{i=1}^{m} v_i$ (, where $v_i = -y_i$,) is the sufficient statistic of the parameter δ for this class of distributions with the probability density function

(2.24) $\qquad f_m(u_m ; \delta)d\mu(u_m) = exp\left\{-\delta u_m - m \cdot \dfrac{\delta^2}{2} - \dfrac{u_m^2}{2m} + log\,(2\pi m)^{-\frac{1}{2}}\right\}d\mu(u_m)$

$$(m = 1, 2, \cdots),$$

where

(2.25) $\qquad b(\delta) = -\dfrac{\delta^2}{2}\,, \qquad a_m(u_m) = -\dfrac{u_m^2}{2m} + log\,(2\pi m)^{-\frac{1}{2}}\,, \quad \text{and} \quad \delta^* = 0\,.$

We assume that the parameter δ has an *a priori* normal distribution with zero mean and known variance $\sigma_1^2 = \sigma_0^2/(2\sigma^2)$. Letting $z = \delta/\sigma_1$, we have

(2.26) $\qquad g(\delta)d\delta = \varphi(z)dz\,, \quad \text{where} \quad \varphi(z) = (2\pi)^{-\frac{1}{2}} \cdot exp\,\{-z^2/2\}\,.$

From (2.10), we have

(2.27) $\qquad\qquad\qquad h(\delta) = \dfrac{\delta_1 + \delta_0 - 2\delta}{\delta_1 - \delta_0}\,.$

Therefore, we get immediately from (2.8′), (2.19) and (2.20)

$$s = -\dfrac{\delta_1 + \delta_0}{2}\,,$$

(2.28) $\qquad\qquad\qquad L(\delta) \sim \dfrac{1}{1 + e^{2(s+\delta)K}}\,,$

$$E_\delta(n) \sim \dfrac{K}{(s+\delta)} \cdot \dfrac{(e^{2(s+\delta)K} - 1)}{(e^{2(s+\delta)K} + 1)}\,.$$

Thus, substituting (2.26) and (2.28) into (2.18), we obtain

(2.29) $\quad \overline{E\,Loss}/NC = -\sigma_1\left\{\displaystyle\int_{-\infty}^{0} z\varphi(z)dz - \dfrac{K}{N}\int_{-\infty}^{\infty} \dfrac{1}{(\sigma_1 z + s)} \cdot \dfrac{(e^{2(\sigma_1 z + s)K} - 1)}{(e^{2(\sigma_1 z + s)K} + 1)} z\varphi(z)dz\right.$

$$\left. -\int_{-\infty}^{\infty} \dfrac{z\varphi(z)}{(e^{2(\sigma_1 z + s)K} + 1)}dz + \dfrac{2K}{N}\int_{-\infty}^{\infty} \dfrac{1}{(\sigma_1 z + s)} \cdot \dfrac{(e^{2(\sigma_1 z + s)K} - 1)}{(e^{2(\sigma_1 z + s)K} + 1)^2} z\varphi(z)dz\right\}\,.$$

Computing (2.29), we could determine the optimum K^* and s^*. But, in the paper [1], T. Colton has discussed the case when $s = 0$, and computed the maximized values of the considered over-all expected net gain at the various given values of $R = N\sigma_1^2$, and finally evaluated the advantage of the sequential optimum plan over the fixed sample size optimum plan.

(2°) **The Fourth Case:** We are secondly concerned with the case when population mean μ_A of a normal population with known variance σ^2 is unknown to us and a mean effect μ_B of treatment B is a known value.

When y_i's $(i = 1, 2, \cdots, m)$ are drawn independently from the normal population with unknown population mean $\delta = (\mu_A - \mu_B)/\sigma$ and unite variance, $u_m = \sum_{i=1}^{m} v_i$ (, where

$v_i = -y_i$,) is the sufficient statistic of the parameter δ for this class of distributions, where $b(\delta) = -\delta^2/2$, $a_m(u_m) = -u_m^2/(2m) + log\,(2\pi m)^{-\frac{1}{2}}$, and $\delta^* = 0$.

We assume that the parameter δ has an *a priori* normal distribution with zero mean and known variance $\sigma_1^2 = \sigma_0^2/\sigma^2$. Letting $z = \delta/\sigma_1$, we have

$$g(\delta)d\delta = \varphi(z)dz\,, \quad \text{where} \quad \varphi(z) = (2\pi)^{-\frac{1}{2}} \cdot exp\,\{-z^2/2\}\,.$$

From (2.19) and (2.20), we have then

(2.30)
$$\begin{cases} L(\delta) \sim \dfrac{1}{e^{2(\sigma_1 z + s)K} + 1}\,, \\[2mm] E_\delta(n) \sim \dfrac{K}{(\sigma_1 z + s)} \cdot \dfrac{(e^{2(\sigma_1 z + s)K} - 1)}{(e^{2(\sigma_1 z + s)K} + 1)}\,. \end{cases}$$

Substituting (2.30) into (2.23), we obtain

(2.31)
$$\overline{E\,Loss}/NC = -\sigma_1 \Big\{ \int_{-\infty}^0 z\varphi(z)dz - \int_{-\infty}^\infty \frac{1}{(e^{2(\sigma_1 z + s)K} + 1)} z\varphi(z)dz$$
$$+ \frac{K}{N} \int_{-\infty}^\infty \frac{K}{(\sigma_1 z + s)} \cdot \frac{(e^{2(\sigma_1 z + s)K} - 1)}{(e^{2(\sigma_1 z + s)K} + 1)^2} z\varphi(z)dz \Big\}\,.$$

Computing (2.31), we could determine the optimum values K^* and s^*. But, in the previous paper [10], we have been concerned with the case when $s = 0$, and computed the maximized values of the over-all expected net gain, and finally evaluated the advantage of the sequential plan over the fixed sample size plan.

Example 2: Binomial Distribution.

(1°) **The Third Case:** In the first place, let us consider the case when population proportions (i. e., effective proportions) $\mathfrak{p}_A$ and $\mathfrak{p}_B$ of the two binomial populations are both unknown to us.

Let us draw $2n$ from N individuals in the form of pair and perform treatment A on one member of the pair and treatment B on the other.

Let a be the outcome of an observation from performing treatment A on one individual and b the outcome of an observation from performing treatment B on the other individual.

Let $a_1, \cdots, a_n$ be the results based on performing treatment A and $b_1, \cdots, b_n$ the results based on performing treatment B, where a_i and b_i are values 1 or 0, and where value 1 denotes that the performed treatment is effective and value 0 denotes that the performed treatment is ineffective $(i = 1, 2, \cdots, n)$.

These results are arranged in the order observed. Thus we can consider the sequence of n pairs:

$$(a_1, b_1), \cdots, (a_n, b_n)\,.$$

Let t_1 be the number of pairs $(1, 0)$ and t_2 the number of pairs $(0, 1)$ in this sequence and let $t = t_1 + t_2$. We consider only the pairs $(0, 1)$ and $(1, 0)$.

Since each pair (a_i, b_i) is equal to one of the pairs $(0, 1)$ and $(1, 0)$, the conditional probability that it is equal to $(0, 1)$ is given by

$$\mathfrak{p} = \frac{(1-\mathfrak{p}_A)\mathfrak{p}_B}{\mathfrak{p}_A(1-\mathfrak{p}_B) + (1-\mathfrak{p}_A)\mathfrak{p}_B}$$

and the conditional probability that it is equal to $(1, 0)$ is given $1-\mathfrak{p}$.

When y_i's $(i = 1, 2, \cdots, t)$ are drawn independently from the binomial population with an unknown population proportion $\mathfrak{p}$, $u_t = \sum_{i=1}^{t} y_i$ is the sufficient statistic of the parameter $\delta = \log \mathfrak{p}^{-1}(1-\mathfrak{p}) = -\log u$ for this class of distributions with the probability density function

$$ f_t(u_t ; \delta) d\mu(u_t) = exp \left\{ -\delta u_t - t \log (1 + e^{-\delta}) - \log [(t+1) \cdot B(t - u_t + 1, u_t + 1)] \right\} d\mu(u_t) $$

$(t = 1, 2, \cdots)$, where $b(\delta) = -\log(1 + e^{-\delta})$, and where u denotes an odds ratio $(\mathfrak{p}_B(1-\mathfrak{p}_A))$ $/(\mathfrak{p}_A(1-\mathfrak{p}_B))$, and where μ is a step-function having jumps of height 1 at every non-negative integer u_t.

Let us define $\delta^* = 0$. We have then that the acceptance of $\delta \in \Omega_0$ is equivalent to the acceptance of the event "$1 < u$" and the acceptance of $\delta \in \Omega_1$ is equivalent to the acceptance of the event "$1 > u$".

Let us assume that the *a priori* distribution for the parameter δ has a uniform distribution with the probability density function $g(\delta) = 1/(\Delta_2 - \Delta_1)$, where $-\infty < \Delta_1 < 0 < \Delta_2 < \infty$.

Letting $e^{-\Delta_i} = d_i$ $(i = 1, 2)$, we have the *a priori* defined interval $\langle d_2, d_1 \rangle$ for the odds ratio u $(0 < d_2 < 1 < d_1 < \infty)$.

We can now compute the over-all expected loss $\overline{E\,Loss}/NC$ of (2.18).

In the previous paper [11], we have been in detail concerned with this case, and then determined the optimum values K^* with respect to various numerical examples for each of fixed several values of s, and consequently computed the minimum values of the considered over-all expected loss function.

(2°) **The Fourth Case:** In the second place, let us consider the case when one population proportion (i. e., effective proportion) $\mathfrak{p}_A$ of the two binomial populations is unknown to us and the other $\mathfrak{p}_B$ known.

When y_i's $(i = 1, 2, \cdots, m)$ are drawn independently from the binomial population with an unknown population proportion $\mathfrak{p}_A$, $u_m = \sum_{i=1}^{m} y_i$ is the sufficient statistic of the parameter $\delta = \log \mathfrak{p}_A^{-1}(1-\mathfrak{p}_A)$ for this class of distributions with the probability density function such that $b(\delta) = -\log(1 + e^{-\delta})$, $a_m(u_m) = -\log((m+1) \cdot B(m - u_m + 1, u_m + 1))$, and $\delta^* = \log \mathfrak{p}_B^{-1}(1-\mathfrak{p}_B)$.

Thus we can compute the over-all expected loss from (2.23), and determine the optimum values K^* and s^*.

In the previous paper [8], we have been in detail concerned with the case when (1) the *a priori* distribution is a uniform distribution with the probability density function $g(\delta) = 1/(\Delta_2 - \Delta_1)$ and (2) $s = \mathfrak{p}_B$. We have then determined the optimum values K^* concerning various given types of *a priori* distribution and consequently computed the minimum values of the considered over-all expected loss and finally evaluated the advantage of the sequential plan over the fixed sample size plan.

Example 3: Gamma-Distribution $\Gamma(m, \delta)$.

Let us consider the case when v_i's $(i = 1, 2, \cdots, m)$ are drawn independently from the Gamma-distribution $\Gamma(1, \delta)$ with unknown mean $1/\delta$, where $0 < v_i < \infty$ and $0 < \delta < \infty$.

Then $u_m = \sum_{i=1}^{m} v_i$ is the sufficient statistic for the parameter δ with the probability density function

(2.31) $f_m(u_m ; \delta) d\mu(u_m) = exp \{-\delta u_m + m \log \delta + \log [\Gamma(m)^{-1} \cdot u_m^{m-1}]\} d\mu(u_m)$,

where

(2.32) $b(\delta) = \log \delta \quad and \quad a_m(u_m) = \log [\Gamma(m)^{-1} \cdot u_m^{m-1}]$,

and where m is any positive integer.

From (2.11) based on **Kitagawa's Theorem**, the function $h(\delta) \doteq 0$ may be computed by solving the following equation.

$$\log [h(\delta) \cdot (\delta_1 - \delta_0) + \delta] = h(\delta) \cdot \{\log (\delta_1 \delta_0^{-1})\} + \log \delta ,$$

that is,

(2.33) $(\delta_1 - \delta_0) \cdot h(\delta)/\delta = (\delta_1 \delta_0^{-1})^{h(\delta)} - 1$,

since $\delta \neq 0$.

Letting

(2.34) $(\delta_1 \delta_0^{-1})^{h(\delta)} = \xi \qquad (, where \ \xi \neq 1 \ and \ \xi > 0)$,

we obtain the following (5.11′) instead of (5.11).

(2.33′) $\{\delta \log (\delta_1 \delta_0^{-1})\}/(\delta_1 - \delta_0) = \log \xi/(\xi - 1)$.

Since the slope $s = \{b(\delta_1) - b(\delta_0)\}/(\delta_1 - \delta_0) = \{\log (\delta_1 \delta_0^{-1})\}/(\delta_1 - \delta_0)$, the relation between δ and ξ is obtained as follows.

(2.35) $\delta = (\log \xi)/\{s(\xi - 1)\}$,

where let s be a non-zero real number.

We also obtain from (2.33), (2.34), and (2.35)

(2.36) $(\delta_0 - \delta_1) \cdot h(\delta) = -\dfrac{\log \xi}{s}$.

In the case of our exponential type, $L(\delta)$ and $E_\delta(n)$ may be obtained as function $L(\xi)$ and $E_\xi(n)$ of ξ respectively from (2.19), (2.20), (2.35), and (2.36), that is,

(2.37) $L(\delta) = L(\xi) \sim \dfrac{\xi^{K/s}}{\xi^{K/s} + 1}$,

(2.38) $E_\delta(n) = E_\xi(n) \sim \dfrac{K}{s} \cdot \dfrac{\xi^{K/s} - 1}{\xi^{K/s} + 1} \cdot \dfrac{\log \xi}{\xi - 1 - \log \xi}$.

It should be noted that both $L(\delta)$, $E_\delta(n)$ are independent with respect to δ_1, δ_0.

Thus, the over-all expected loss of the present case results as follows, from (2.18), (2.35), (2.37), (2.38), and our assumption,

(2.39) $\overline{E \, Loss}/NC = \displaystyle\int_{\Delta_1}^{\delta^*} (\delta^* - \delta) g(\delta) d\delta - \dfrac{1}{N} \int_{\Delta_1}^{\Delta_2} (\delta^* - \delta) E_\delta(n) g(\delta) d\delta$

$$- \int_{\Delta_1}^{\Delta_2} (\delta^* - \delta) L(\delta) g(\delta) d\delta + \dfrac{2}{N} \int_{\Delta_1}^{\Delta_2} (\delta^* - \delta) E_\delta(n) L(\delta) g(\delta) d\delta ,$$

where $g(\delta)$ denotes the probability density function of the *a priori* distribution for δ.

From (2.35), we have

(2.39′) $$\overline{E\,Loss}/NC = \phi_1(K,\,s\,|\,\xi_1,\,\xi^*) - \frac{1}{N}\,\phi_2(K,\,s\,|\,\xi_1,\,\xi_2)$$

$$-\phi_3(K,\,s\,|\,\xi_1,\,\xi_2) + \frac{2}{N}\,\phi_4(K,\,s\,|\,\xi_1,\,\xi_2)\,,$$

where

$$\phi_1(K,\,s\,|\,\xi_1,\,\xi^*) = \frac{1}{s}\int_{\xi_1}^{\xi^*}\Big(\delta^* - \frac{\log\xi}{s(\xi-1)}\Big)\cdot\frac{(\xi-1-\xi\,\log\xi)}{\xi(\xi-1)^2}\,g_1(\xi)d\xi\,,$$

$$\phi_2(K,\,s\,|\,\xi_1,\,\xi_2) = \frac{K}{s^2}\int_{\xi_1}^{\xi_2}\Big(\delta^* - \frac{\log\xi}{s(\xi-1)}\Big)\cdot\frac{(\xi-1-\xi\,\log\xi)\,\log\xi}{(\xi-1-\log\xi)\xi(\xi-1)^2}\cdot\frac{\xi^{K/s}-1}{\xi^{K/s}+1}\,g_1(\xi)d\xi\,,$$

$$\phi_3(K,\,s\,|\,\xi_1,\,\xi_2) = \frac{1}{s}\int_{\xi_1}^{\xi_2}\Big(\delta^* - \frac{\log\xi}{s(\xi-1)}\Big)\cdot\frac{(\xi-1-\xi\,\log\xi)}{\xi(\xi-1)^2}\cdot\frac{\xi^{K/s}}{\xi^{K/s}+1}\,g_1(\xi)d\xi\,,$$

$$\phi_4(K,\,s\,|\,\xi_1,\,\xi_2) = \frac{K}{s^2}\int_{\xi_1}^{\xi_2}\Big(\delta^* - \frac{\log\xi}{s(\xi-1)}\Big)\cdot\frac{(\xi-1-\xi\,\log\xi)\,\log\xi}{(\xi-1-\log\xi)\xi(\xi-1)^2}\cdot\frac{\xi^{K/s}(\xi^{K/s}-1)}{(\xi^{K/s}+1)^2}\,g_1(\xi)d\xi\,,$$

and where

$$\varDelta_i = (\log\xi_i)/\{s(\xi_i-1)\}\quad(i=1,\,2)\,,\qquad\delta^* = (\log\xi^*)/\{s(\xi^*-1)\}\,,$$

and

$$g(\delta) = g_1(\xi)\,.$$

Thus we must determine the optimum values K^* and s^* of K and s respectively so that (2.39′) is minimized.

Example 4: Poisson-Distribution $P(\lambda)$.

Let us consider the case when v_i's $(i=1,\,2,\,\cdots,\,m)$ are drawn independently from the Poisson distribution with an unknown mean λ, $0<\lambda<\infty$. Then $u_m = \sum_{i=1}^{m} v_i$ is a sufficient statistic for the parameter $\delta = -\log\lambda$ with the probability density function

(2.40) $$f_m(u_m\,;\,\delta)d\mu(u_m) = exp\,\{-u_m\delta - me^{-\delta} + \log\,[m^{u_m}\cdot\varGamma(u_m+1)^{-1}]\}d\mu(u_m)\,,$$

where

(2.41) $$b(\delta) = -e^{-\delta}\quad\text{and}\quad a_m(u_m) = \log\,[m^{u_m}\cdot\varGamma(u_m+1)^{-1}]\,,$$

and where μ is a step-function having jumps of height 1 at every nonnegative integer u_m. Moreover, we assume that an *a priori* defined interval for the parameter δ is $\langle\varDelta_1,\,\varDelta_2\rangle\,(-\infty<\varDelta_1\leqq\delta\leqq\varDelta_2<\infty)$.

From (2.10) based on **Kitagawa's Theorem**, the function $h(\delta)\neq0$ may be computed by solving the following equation.

(2.42) $$exp\,\{-(\delta_1-\delta_0)h(\delta)\} - 1 = h(\delta)\cdot exp\,\{\delta\}\cdot(exp\,\{-\delta_1\} - exp\,\{-\delta_0\})\,.$$

Letting

(2.43) $$exp\,\{-(\delta_1-\delta_0)h(\delta)\} = \xi\qquad(\,,\,\text{where}\,\,\xi\neq1\,\,\text{and}\,\,\xi>0)\,,$$

we obtain the following (2.42′) instead of (2.42).

(2.42′) $$-\frac{(\delta_1-\delta_0)e^{-\delta}}{e^{-\delta_1}-e^{-\delta_0}} = \frac{\log\xi}{\xi-1}\,.$$

Since the slope $s = \{b(\delta_1)-b(\delta_0)\}/(\delta_1-\delta_0) = \{-e^{-\delta_1}+e^{-\delta_0}\}/(\delta_1-\delta_0)$, the relation between δ and ξ is obtained.

(2.44)
$$e^{-\delta} = \frac{s \log \xi}{\xi - 1},$$

where let s be a nonzero real number.

Also, from (2.43), we obtain

(2.45)
$$(\delta_0 - \delta_1) \cdot h(\delta) = \log \xi.$$

Thus, in the case of our Poisson type, $L(\delta)$ and $E_{\delta}(n)$ may be obtained as functions $L(\xi)$ and $E_{\xi}(n)$ of ξ respectively from (2.19), (2.20), and (2.45), that is,

(2.46)
$$L(\delta) = L(\xi) \sim \frac{1}{1 + \xi^K},$$

(2.47)
$$E_{\delta}(n) = E_{\xi}(n) \sim \frac{K}{s} \cdot \frac{\xi - 1}{(\log \xi) + 1 - \xi} \cdot \frac{1 - \xi^K}{1 + \xi^K},$$

where let us note $b'(\delta) = e^{-\delta} = (s \log \xi)/(\xi - 1)$ from (2.44). It should be noted that both $L(\delta)$ and $E_{\delta}(n)$ are independent with respect to δ_0, δ_1.

Thus, the over-all expected loss of the present case results as follows, from (2.18), (2.44), (2.46), (2.47), and our assumption.

(2.48)
$$\overline{E\ Loss}/NC = \psi_1(K, s \,|\, \xi_1, \xi^*) - \frac{1}{N} \psi_2(K, s \,|\, \xi_1, \xi_2)$$

$$-\psi_3(K, s \,|\, \xi_1, \xi_2) + \frac{2}{N} \psi_4(K, s \,|\, \xi_1, \xi_2),$$

where

$$\psi_1(K, s \,|\, \xi_1, \xi^*) = -\int_{\xi_1}^{\xi^*} \Big(\delta^* - \log \frac{(\xi - 1)}{s \log \xi}\Big) \cdot \frac{(\xi - 1 - \xi \log \xi)}{\xi(\xi - 1) \log \xi} g_1(\xi) d\xi,$$

$$\psi_2(K, s \,|\, \xi_1, \xi_2) = -\frac{K}{s} \int_{\xi_1}^{\xi_2} \Big(\delta^* - \log \frac{(\xi - 1)}{s \log \xi}\Big) \cdot \frac{(\xi - 1 - \xi \log \xi)}{(1 - \xi + \log \xi)\xi \log \xi} \cdot \frac{(1 - \xi^K)}{(1 + \xi^K)} g_1(\xi) d\xi,$$

$$\psi_3(K, s \,|\, \xi_1, \xi_2) = -\int_{\xi_1}^{\xi_2} \Big(\delta^* - \log \frac{(\xi - 1)}{s \log \xi}\Big) \cdot \frac{(\xi - 1 - \xi \log \xi)}{\xi(\xi - 1) \log \xi} \cdot \frac{1}{(1 + \xi^K)} g_1(\xi) d\xi,$$

$$\psi_4(K, s \,|\, \xi_1, \xi_2) = -\frac{K}{s} \int_{\xi_1}^{\xi_2} \Big(\delta^* - \log \frac{(\xi - 1)}{s \log \xi}\Big) \cdot \frac{(\xi - 1 - \xi \log \xi)}{(1 - \xi + \log \xi)\xi \log \xi} \cdot \frac{(1 - \xi^K)}{(1 + \xi^K)^2} g_1(\xi) d\xi,$$

and where

$$e^{-\Delta i} = (s \log \xi_i)/(\xi_i - 1) \quad (i = 1, 2), \qquad e^{-\delta^*} = (s \log \xi^*)/(\xi^* - 1),$$

and

$$g(\delta) = g_1(\xi).$$

Thus we must determine the optimum values K^* and s^* of K and s respectively so that $\overline{E\ Loss}/NC$ of (2.48) is minimized.

Acknowledgement

Reference

[1] Colton, T., "A model for Selecting One of Two Medical Treatments", *J. Am. Stat. Assoc.*, **58**, 388–400 (1963),

[2] Colton, T., "A Two-Stage Model for Selecting One of Two Treatments", *Biometrics*, **21**, 169–180 (1965).

[3] Kitagawa, T., "Successive Process of Statistical Inference Associated with an Additive Family of Sufficient Statistics", *Bull. Math. Stat.*, **7**, No. 3~4, 92–112 (1957).

[4] Sugimura, M., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 1", *Mem. Fac. Gen. Educ. Kumamoto Univ.*, *Ser. Nat. Sci.*, **2**, 1–7 (1967).

[5] Sugimura, M., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 2", *Kumamoto J. Sci. Ser. A*, **7**, No. 4, 87–94 (1967).

[6] Sugimura, M., and Asano, C., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 3", *Kumamoto J. Sci. Ser. A*, **7**, No. 4, 95–102 (1967).

[7] Sugimura, M., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 4.1", *J. Kumamoto Women's Univ.*, **19**, 65–71 (1967).

[8] Sugimura, M., and Asano, C., "Optimum Designs for Selecting One of Two Medical Treatments, Sequential Plan 1", *Bull. Math. Stat.*, **12**, No. 3~4, 1–9 (1967).

[9] Sugimura, M., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 5", *J. Kumamoto Women's Univ.*, **20**, (in press, 1968).

[10] Sugimura, M., "Optimum Designs for Selecting One of Two Treatments, Sequential Plan 2", *Kumamoto J. Sci. Ser. A*, **8**, No. 1, 11–19 (1967).

[11] Sugimura, M., and Asano, C., "Optimum Designs for Selecting One of Two Treatments, Fixed Sample Size Plan 6 and Sequential Plan 3", *Kumamoto J. Sci. Ser. A*, **8**, No. 1, 21–57 (1967).

[12] Sugimura, M., "Optimum Designs for Selecting One of Two Treatments, General Considerations", *Kumamoto J. Sci.*, *Ser. A*, **8**, No. 2 (in press, 1967).

[13] Wald, A., "Sequential Analysis", John Wiley and Sons: New York (1947).