Kumamoto J. Sci., Ser. A, Vol. 8, No. 1, pp. 11-19 (1967)

# OPTIMUM DESIGNS FOR SELECTING ONE OF TWO TREATMENTS, SEQUENTIAL PLAN 2

By

Masahiko SUGIMURA

(Received April 1, 1967)

## 1. Introduction.

Our present question is how it is possible to design an optimum clinical trial when a total of patients with a disease are to be treated with one of the two medical treatments, where the therapeutic efficacy is unknown for one of the two treatments.

In the planning of medical experiments to assess the therapeutic efficacy of new drugs or treatments, a most important question how large to make the clinical trial. On the one hand, experimenter wants as few patients as possible to participate so that the number of patients receiving the inferior treatment during all of the clinical trial is minimized, so that the clinical trial is brought to a conclusion as speedily as possible and so that the results may be quickly utilized to treat the many remaining patients with the disease in question. On the other hand, a sufficient number of patients must participate so that experimenter can reasonably make sure that the truly superior treatment is selected and consequently administered on as many patients with the disease in question as possible.

In the situation of this kind, an application of Neyman-Pearson principle will lose its active meaning. As an alternative, it seems reasonable to approach the problem from the point of view of the consequences of decisions maid. Therefore, we should like to introduce the concepts of the loss and the gain.

In our medical setting, the consequences of both right and wrong decisions are the principal concern from an ethical point of view. Whence, when we consider the loss or the gain, it will be desirable to disregard all others such as the costs of experimentation and confine ourselves to the loss or the gain due to the consequences of treating a patient with the superior or the inferior of the two treatments.

In this paper, the sequential plan will be proposed for the model in normal population. Namely, a clinical trial will be sequentially performed on each patient chosen at random one by one from a group of patients by administering the one treatmeant of which the therapeutic efficacy is unknown.

In the next place, the effects of the two treatments will be compared by observing the responses based on performing the particular treatment from the first patient until $n$-th on each patient and a decision for selecting one of the two treatments will be finally determined as the result. Thus, the treatment that is to be selected as the better at the conclusion of the beforementioned clinical trial will be performed on the all remaining patients.

The problem therefore is to determine the optimum procedure containing the optimum location of the boundaries so that the loss constructed on the basis of the proposed procedure is minimized or the gain is maximized.

The sequential plan has already been investigated by T. Colton [4] in the case of the two normal populations with unknown means and with known and common variances. The case of the two populations will be discussed throughout this paper where one population mean is unknown and the other known. Finally, a comparison will be numerically made of the optimal sequential plan to the optimal fixed sample size plan [5].

## 2. Assumptions.

The next several assumptions are prepared throughout this paper as follows. These are

(1) There are $N$ patients with a disease who are to be treated with one of the two medical treatments, which are unknown on their superiority or inferiority, denoted by $A$ and $B$. $N$ is fixed and large.

(2) It is assumed that we obtain a quantitative measure of the preassigned response for each individual. The population mean, denoted by $\mu_A$, of individual responses due to treatment $A$ is known. On the other hand, the individual response due to treatment $B$ is a normally distributed variate $X_B$ with unknown mean $\mu_B$ and known variance $\sigma^2$. Moreover, we shall assume that higher response is associated with better effect.

Letting $\mu_A-\mu_B=\delta$, we should then like to retain treatment $A$ for the trial if $\delta$ is positive and select treatment $B$ instead of treatment $A$ if $\delta$ is negative.

(3) The only cost involved is the consequence of treating a patient with the superior or the inferior of the two treatments, and all other costs may be and will be regarded from our ethical point of view. Namely, the cost is directly proportional to the true mean difference $\delta$ between the two means. Based on the loss formulation one can state that if a patient is treated with the inferior treatment a positive loss proportional to the true mean difference $\delta$ is scored for him. Whereas, based on the net gain formulation one can state

that if a patient is treated with the superior treatment a positive gain proportional to the true mean difference $\delta$ is scored for him, while if treated with the inferior treatment a negative gain proportional to the true difference $\delta$ is scored for him.

(4) Moreover, we shall assume that the true mean $\mu_B$ of treatment $B$ is distributed in accordance with an *a priori* normal distribution with mean $\mu_A$ and known variance $\sigma_0^2$.

### 3. Procedure.

The trial no longer calls for a fixed number of participants, but the trial is sequentially performed on each patient by administering treatment $B$. After the results from each patient are available, a decision based on the cumulative evidence is made to select one of the two treatments as the better and use it on all remaining patients, or to continue the trial by having an additional participate.

After $n$-th stage, compute the cumulative sum $d=\sum_{i=1}^{n} x_i$ of $n$ observations obtained by performing treatment $B$ from the first patient until the $n$-th.

Procedure:

If $d \geqq K\sigma + n\mu_A$, use treatment $B$ on the remaining $N-n$ patients;

If $d \leqq -K\sigma + n\mu_A$, use treatment $B$ on the remaining $N-n$ patients;

If $-K\sigma + n\mu_A < d < K\sigma + n\mu_A$, continue treatment $B$ on an additional patients, where $K$ is a parameter that denotes the location of the boundaries between an acceptance line and a rejection line.

We should like to determine an optimum value $K^*$ of $K$ so that the below gain is maximized.

### 4. Construction of Overall Expected Loss Function.

We shall assume that a decision of selecting one of the two treatments has now been made in the $n$-th stage, and denote by $L(\mu_B)$ the probability of selecting treatment $A$ as the result of the decision when $\mu_B$ is the true mean value of the observations based on performing treatment $B$.

From the procedure proposed in the above, $L(\mu_B)$ is expressed in the next approximate formula.

$$(1) \qquad L(\mu_B)=\frac{e^{2K(\mu_A-\mu_B)/\sigma}}{e^{2K(\mu_A-\mu_B)/\sigma}+1}=\frac{e^{az}}{e^{az}+1},$$

where $a=2\sigma_0 K/\sigma$, $z=\delta/\sigma_0$.

Whence, the probability of selecting treatment $B$ as the result of the decision is

$$1-L(\mu_B)=\frac{1}{e^{az}+1}. \tag{1'}$$

Moreover, the following approximation formula is derived for the expected value $E\mu_B(n)$ of the number $n$ of observations required by the decision plan.

$$E\mu_B(n)=\frac{K\sigma(e^{az}-1)}{\sigma_0 z(e^{az}+1)}, \tag{2}$$

where $a=2\sigma_0 K/\sigma$, $z=\delta/\sigma_0$.

If treatment $B$ is inferior to $A$ (i.e., $\delta>0$), then the expected loss $[Loss]_B$ based on performing treatment $B$ results from assumption (3) as follows.

$$\begin{aligned}[Loss]_B &= C\delta[E(n)+\{N-E(n)\}Pr(\text{Selecting } B)] \\ &= C\delta[E(n)+\{N-E(n)\}\{1-L(\mu_B)\}] \\ &= NC\sigma_0\left[\frac{ae^{az}(e^{az}-1)}{4R(e^{az}+1)^2}+\frac{z}{e^{az}+1}\right],\end{aligned} \tag{3}$$

where $E(n)=E\mu_B(n)$, $R=N\sigma_0^2/2\sigma^2$ and $C$ is a proportionality factor.

On the other hand, if treatment $A$ is inferior to $B$ (i.e., $\delta<0$), then the expected loss $[Loss]_A$ based on performing treatment $A$ results from assumption (3) as follows.

$$\begin{aligned}[Loss]_A &= -C\delta\{N-E(n)\}Pr(\text{Selecting } A) \\ &= NC\sigma_0\left[\frac{ae^{az}(e^{az}-1)}{4R(e^{az}+1)^2}-\frac{ze^{az}}{e^{az}+1}\right].\end{aligned} \tag{4}$$

We shall now obtain the overall expected loss $\overline{E\ Loss}$ (5) by averaging (3) and (4) over an *a priori* distribution for $\delta$ (, that is, for $z$).

$$\begin{aligned}\overline{E\ Loss}/NC\sigma_0 &= \int_{-\infty}^{0}[Loss]_A\varphi(z)dz+\int_0^{\infty}[Loss]_B\varphi(z)dz \\ &= \frac{a}{4R}\int_{-\infty}^{0}\frac{e^{az}(e^{az}-1)}{(e^{az}+1)^2}\varphi(z)dz-\int_{-\infty}^{0}\frac{e^{az}}{e^{az}+1}z\varphi(z)dz \\ &\quad+\frac{a}{4R}\int_0^{\infty}\frac{e^{az}(e^{az}-1)}{(e^{az}+1)^2}\varphi(z)dz+\int_0^{\infty}\frac{1}{e^{az}+1}z\varphi(z)dz \\ &= \frac{a}{4R}\int_{-\infty}^{\infty}\frac{e^{az}(e^{az}-1)}{(e^{az}+1)^2}\varphi(z)dz+2\int_0^{\infty}\frac{1}{e^{az}+1}z\varphi(z)dz \\ &= \frac{a}{8R}\int_{-\infty}^{\infty}\left(\frac{e^{az}-1}{e^{az}+1}\right)^2\varphi(z)dz\end{aligned} \tag{5}$$

$$+\frac{a}{8R}\int_{-\infty}^{\infty}\frac{e^{az}-1}{e^{az}+1}\varphi(z)dz+2\int_{0}^{\infty}\frac{1}{e^{az}+1}z\varphi(z)dz$$

$$=\frac{a}{4R}\int_{0}^{\infty}\left(\frac{e^{az}-1}{e^{az}+1}\right)^{2}\varphi(z)dz+2\int_{0}^{\infty}\frac{1}{e^{az}+1}z\varphi(z)dz,$$

where $\varphi(z)=(2\pi)^{-\frac{1}{2}}\exp(-z^2/2)$.

## 5. Construction of Expected Net Gain Function.

We shall assume as before that a decision of selecting one of the two treatments has now been made in the $n$-th stage.

Whether treatment $A$ is superior to treatment $B$ or $B$ superior to $A$, the both gain functions have the same formulas. Whence, in point of constructing the gain function based on assumption (3) and the procedure, it may be and will be assumed without loss of generality that $\delta$ is positive. Namely, the expected net gain function *E Net Gain* is obtained as follows.

(6) $$E\ Net\ Gain=G\delta\left[-E(n)+\{N-E(n)\}\{Pr(\text{Selecting } A)-Pr(\text{Selecting } B)\}\right]$$

$$=NG\sigma_0\left[-\frac{a(e^{az}-1)}{4R(e^{az}+1)}-\frac{a}{4R}\left(\frac{e^{az}-1}{e^{az}+1}\right)^{2}+z\cdot\frac{e^{az}-1}{e^{az}+1}\right],$$

where $z=\delta/\sigma_0$, $a=(2\sigma_0K)/\sigma$, and $R=(N\sigma_0^2)/(2\sigma^2)$ as before.

We shall now obtain the overall expected net gain $\overline{E\ Net\ Gain}$ (7) by averaging (6) over an *a priori* distribution for $\delta$ (, that is, for $z$).

(7) $$\overline{E\ Net\ Gain}/NG\sigma_0=-\frac{a}{4R}\int_{-\infty}^{\infty}\frac{e^{az}-1}{e^{az}+1}\varphi(z)dz$$

$$-\frac{a}{4R}\int_{-\infty}^{\infty}\left(\frac{e^{az}-1}{e^{az}+1}\right)^{2}\varphi(z)dz+\int_{-\infty}^{\infty}\frac{e^{az}-1}{e^{az}+1}z\varphi(z)dz.$$

Moreover,

$$\int_{-\infty}^{\infty}\frac{e^{az}-1}{e^{az}+1}\varphi(z)dz=0$$

and also

$$\int_{-\infty}^{\infty}\left(\frac{e^{az}-1}{e^{az}+1}\right)^{2}\varphi(z)dz=2\int_{0}^{\infty}\left(\frac{e^{az}-1}{e^{az}+1}\right)^{2}\varphi(z)dz,\qquad\text{and}$$

$$\int_{-\infty}^{\infty}\frac{e^{az}-1}{e^{az}+1}\cdot z\varphi(z)dz=2\int_{0}^{\infty}\frac{e^{az}-1}{e^{az}+1}\cdot z\varphi(z)dz.$$

whence, $\overline{E\ Net\ Gain}/NG\sigma_0$ results simply as follows.

$$(7') \qquad \overline{E\ Net\ Gain}/NG\sigma_0 = -\frac{a}{2R}\int_0^\infty \left(\frac{e^{az}-1}{e^{az}+1}\right)^2 \varphi(z)dz + 2\int_0^\infty \frac{e^{az}-1}{e^{az}+1} z\varphi(z)dz .$$

From (5) and (7'), a relation between $\overline{E\ Net\ Gain}/NG\sigma_0$ and $\overline{E\ Loss}/NC\sigma_0$ i easily obtained.

$$(8) \qquad \overline{E\ Net\ Gain}/NG\sigma_0 = -2\,\overline{E\ Loss}/NC\sigma_0 + 2\int_0^\infty z\varphi(z)dz$$

$$= 2\{(2\pi)^{-\frac{1}{2}} - \overline{E\ Loss}/NC\sigma_0\} .$$

By noting that

$$\int_0^\infty \left(\frac{e^{az}-1}{e^{az}+1}\right)^2 \varphi(z)dz = \int_0^\infty \varphi(z)dz - 4\int_0^\infty \frac{e^{az}}{(e^{az}+1)^2}\varphi(z)dz$$

$$= \frac{1}{2} - \frac{2}{a}\int_0^\infty \frac{e^{az}-1}{e^{az}+1} z\varphi(z)dz ,$$

we obtain

$$\overline{E\ Net\ Gain}/NG\sigma_0 = \left(\frac{1}{R}+2\right)\int_0^\infty \frac{e^{az}-1}{e^{az}+1} z\varphi(z)dz - \frac{a}{4R} .$$

Integrating by parts gives

$$(9) \qquad \overline{E\ Net\ Gain}/NG\sigma_0 = 2\left(2+\frac{1}{R}\right)\int_0^\infty \frac{ae^{az}}{(e^{az}+1)^2}\varphi(z)dz - \frac{a}{4R} .$$

We want to determine the $a=(2\sigma_0 K)/\sigma$ which maximizes (9).

Differentiating with respect to $a$ and setting the derivative equal to zero give

$$\frac{1}{8} = (2R+1)\frac{d}{da}\int_0^\infty \frac{ae^{az}}{(e^{az}+1)^2}\varphi(z)dz ,$$

as an equation involving $a$ and $R$. Analytical solution of this equation for $a$ i terms of $R$ is not feasible. Solving for $R$ in terms of $a$ gives

$$(10) \qquad R = \frac{1}{16\dfrac{d}{da}\displaystyle\int_0^\infty ae^{az}(e^{az}+1)^{-2}\varphi(z)dz} - \frac{1}{2} .$$

Thus for arbitrary choice of $a$, (10) gives the $R$ such that the chosen $a$ is op timal.

By substituting (10) into (9), we can express $\overline{E\ Net\ Gain}/NG\sigma_0$ in terms of only, that is,

$$(11)\qquad \overline{E\ Net\ Gain}/NG\sigma_0=4\cdot\frac{\int_0^\infty ae^{az}(e^{az}+1)^{-2}\varphi(z)dz-a\frac{d}{da}\int_0^\infty ae^{az}(e^{az}+1)^{-2}\varphi(z)dz}{1-8\frac{d}{da}\int_0^\infty ae^{az}(e^{az}+1)^{-2}\varphi(z)dz}.$$

For the purpose of obtaining numerical results the integrals must be evaluated. But, since formal integration does not appear feasible, we should like to express $(1+e^{az})^{-2}$ in an infinite series involving terms $e^{-jaz}$ and to then integrate term by term. The results are

$$(12)\qquad \int_0^\infty \frac{ae^{az}}{(1+e^{az})^2}\varphi(z)dz=(2\pi)^{-\frac{1}{2}}\{\sum_{j=1}^{\infty}(-1)^{j+1}jaG(ja)\},$$

where $G(u)$ denotes Mill's Ratio $\Phi(-u)/\varphi(u)$, and by straightforward differentiation,

$$(13)\qquad a\frac{d}{da}\int_0^\infty \frac{ae^{az}}{(1+e^{az})^2}\varphi(z)dz=(2\pi)^{-\frac{1}{2}}\sum_{j=1}^{\infty}(-1)^{j+1}ja[G(ja)+jaG'(ja)].$$

Substituting (12) and (13) into (10) and (11) gives

$$(14)\qquad R=\frac{(2\pi)^{\frac{1}{2}}a}{16\sum_{j=1}^{\infty}(-1)^{j+1}ja[G(ja)+jaG'(ja)]}-\frac{1}{2}$$

and

$$(15)\qquad \overline{E\ Net\ Gain}/NG\sigma_0=4(2\pi)^{-\frac{1}{2}}a\cdot\frac{\sum_{j=1}^{\infty}(-1)^{j}(ja)^2G'(ja)}{a-8(2\pi)^{-\frac{1}{2}}\sum_{j=1}^{\infty}(-1)^{j+1}ja[G(ja)+jaG'(ja)]}.$$

We can now numerically evaluate $\overline{E\ Net\ Gain}/NG\sigma_0$. Arbitrarily, the values of $a$ of 0.5 (0.5) 5.0 were selected. By using (14), $R$ was determined such that the chosen $a$ is optimal. The resulting values of $R$ are shown in the second column of Table 1. The value of $(2\pi)^{\frac{1}{2}}\overline{E\ Net\ Gain}/(2NG\sigma_0)$ when the chosen $a$ is optimal

**Table 1. The Optimum Sequential Plan.**

| $a^*$ | $R$ | $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}[\overline{E\ Net\ Gain}]_{a^*}$ |
|---|---|---|
| 0.5 | 0.09 | 0.201 |
| 1.0 | 0.37 | 0.370 |
| 1.5 | 0.84 | 0.501 |
| 2.0 | 1.57 | 0.601 |
| 2.5 | 2.59 | 0.678 |
| 3.0 | 3.99 | 0.736 |
| 3.5 | 5.82 | 0.781 |
| 4.0 | 8.16 | 0.817 |
| 4.5 | 11.07 | 0.845 |
| 5.0 | 14.62 | 0.867 |

($a^*$ denotes the optimum value of $a$ for given $R$)

is calculated from (15) (see the third column of Table 1).

Even if the expression (15) of $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ when the chosen $a$ is optimal is equal to that of the case of two samples due to Colton's [4], it should be noted that the $a=(2\sigma_0 K)/\sigma$ is equal to $(2a_0)/\sigma$ where $a_0=K\sigma_0$ is a constant $a$ in the expression of $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ due to Colton's [4].

## 6. Comparison of Optimum Fixed Plan and Optimum Sequential Plan.

Finally, we should like to compare the optimal fixed sample size plan in the previous paper [5] with the optimal sequential plan. Of course, intuitively, one expects better results with a sequental plan than a fixed plan, but how much better is actually the sequential plan?

The $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ in the previous paper [5] is compared with the $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ of (15) in this paper. Since both $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ depend on $R$, the comparison also depends on $R$.

Table 2 gives some numerical results of this comparison. The first three columns of Table 2 are a transcription from Table 1. They give the optimal value of the parpmeter $a$ (denoted by $a^*$) of a sequential plan for given $R$ and the optimal value of the $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ of a sequential plan for each $a^*$.

The fourth and fifth columns of the table give the optimal value of the parameter $p$ (denoted by $p^*$) and the optimal value of the $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ for each $p^*$ of a fixed sample size plan, for the same $R$. The last column of the table gives the ratio of $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ in a fixed sample size plan to that in a sequential plan.

The results show that the optimal sequential plan has its greatest advantage over the optimal fixed plan at $R=0.09$. Here the overall expected net gain of the former is 27.3 per cent more than that of the latter. As $R$ increases the relative

**Table 2. Comparison of Optimal Fixed Plan and Optimal Sequential Plan.**

| $R$ | Optimal Sequential Plan | | Optimal Fixed Plan | | Per Cent Additional Gain of Sequential over Fixed |
|---|---|---|---|---|---|
| | $a^*$ | $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ | $p^*$ | $(2\pi)^{\frac{1}{2}}(2NG\sigma_0)^{-1}(\overline{E\ Net\ Gain})$ | |
| 0.09 | 0.5 | 0.201 | 0.321 | 0.158 | 27.3 |
| 0.37 | 1.0 | 0.370 | 0.291 | 0.298 | 24.1 |
| 0.84 | 1.5 | 0.501 | 0.258 | 0.408 | 22.9 |
| 1.57 | 2.0 | 0.601 | 0.226 | 0.498 | 20.6 |
| 2.59 | 2.5 | 0.678 | 0.198 | 0.571 | 18.8 |
| 3.99 | 3.0 | 0.736 | 0.173 | 0.630 | 16.9 |
| 5.82 | 3.5 | 0.781 | 0.153 | 0.678 | 15.4 |
| 8.16 | 4.0 | 0.817 | 0.135 | 0.716 | 14.2 |
| 11.07 | 4.5 | 0.845 | 0.120 | 0.750 | 12.7 |
| 14.62 | 5.0 | 0.867 | 0.108 | 0.782 | 10.8 |

advantage of the sequential plan over the fixed plan decreases. For $R$ as high as 14.62 the optimal sequential plan has an overall expected net gain which is only 10.8 per cent more than that of the optimal fixed plan.

Thus, the dominancy of the sequential plan over the fixed plan may be clearly established numerically.

**Acknowledgement**

The author wishes to express his heartiest thanks to Professor Dr. T. Kitagawa of Kyushu University, for his kind advice and valuable suggestions in connection with this work.

**References**

[1] Anscombe, F.J., "*Sequential Medical Trials*", J. Am. Stat. Assoc., **58**, 365-383 (1963)
[2] Armitage, P., "*Sequential Medical Trials*", Blackwell Scientific Publications: Oxford, England (1961)
[3] Armitage, P., "*Sequential Medical Trials: Some Comments on F.J. Anscombe's Paper*", J. Am. Stat. Assoc., **58**, 384-387 (1963)
[4] Colton, T., "*A Model for Selecting One of Two Medical Treatments*" J. Am. Stat. Assoc., **58**, 388-400 (1963)
[5] Sugimura, M., "*Optimum Designs for Selecting One of Two Treatments, Fixed Sample Size Plan* 5", J. Kumamoto Women's Univ., **20** (1968, to be published)
[6] Wald, A., "*Sequential Analysis*", John Wiley and Sons: New York (1947).

***Department of Mathematics and Physics,***
***Kumamoto Women's University***